Social Media, Social Networks, Stereotypes, and Sustainability

Essays in Behavioral Economics

Dissertation
zur Erlangung des Doktorgrades
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen

vorgelegt von
Raphael Moritz
aus Bergisch Gladbach

Tübingen
2025

1. Betreuerin:	Prof. Dr. Kerstin Pull
2. Betreuer:	Prof. Dr. Patrick Kampkötter
Tag der mündlichen Prüfung:	23.10.2025
Dekanin/Dekan:	Prof. Dr. Taiga Brahm & Prof. Dr. Dominik Papies
1. Gutachterin:	Prof. Dr. Kerstin Pull
2. Gutachter:	Prof. Dr. Stefan Mayer

Overview

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In an era where social media and online (matching) platforms play a substantial role in social and economic life (Allcott et al., 2020; Belleflamme & Peitz, 2021; Cheng et al., 2024; Datareportal, 2024; Edelman et al., 2017; Levy, 2021), the way individuals and organizations signal information has significant implications for both personal and professional outcomes (Spence, 1973, 2002). Whether it is a student searching for a roommate (Moritz & Manger, 2022; Sawert, 2020), a social media user seeking to expand her network (Ajzenman et al., 2025; Evsyukova et al., 2025), or a board of directors setting performance metrics for the firm's top executives (Cohen, Kadach, Ormazabal, & Reichelstein, 2023; Tsang et al., 2021), the availability and interpretation of *non-traditional* information can profoundly affect decisions and shape broader patterns of inequality.

Recent advances in technology allow individuals to curate their online presence and enable firms to incorporate new types of performance measures, such as environmental, social, and governance (ESG) factors, into compensation structures (Ikram et al., 2023; Willis Towers Watson, 2023). Yet, despite this unprecedented capacity to gather and share information, key questions remain: How do these new forms of information affect decision-making processes of individuals in markets and organizations? Can technology-driven transparency reduce biases, or does it merely replicate discrimination in new forms? And to what extent can novel performance metrics actually transform corporate behavior rather than serve as "window dressing"?

This thesis addresses these questions in five chapters. Four of them (Chapters 2 to 5) focus on discrimination and decision-making in informal markets. These studies leverage carefully designed randomized field experiments in which fictitious social media profiles, names, and other cues serve to signal individual traits or identities. By randomly varying certain signals, such as ethnicity, visual stereotypes, or personality cues, these chapters assess the extent to which social media-derived information shapes bias, acceptance rates, and callback decisions. Thereby, this thesis illuminates how small, strategically provided pieces of social media information can mitigate discrimination or reinforce it.

Chapter 6 shifts the perspective to the corporate sphere, using a detailed hand collected panel dataset to examine how ESG performance metrics are integrated into executive compensation contracts. Although analyzing a different question and using a different methodology than in Chapters 2 to 5, this chapter shares a core focus on how *non-traditional* information and signals, such as non-financial sustainability metrics, affect market outcomes. Across both the interpersonal and corporate settings, this thesis demonstrates how new information channels and incentive mechanisms can (re-)shape market outcomes, though not always in the ways one might expect.

Following the general introduction (Chapter 1), this thesis comprises three essays on discrimination, information, and social media (Chapters 2 to 4), followed by a fourth essay on personality and social media (Chapter 5), and a fifth essay on ESG metrics in executive compensation (Chapter 6). Chapter 7 concludes. In what follows, I provide a concise introduction of each thematic focus, concluding with a short summary.

Discrimination, Information, and Social Media Discrimination remains pervasive and widespread across multiple social and economic domains. Members of ethnic minority groups have lower average incomes, have a harder time finding jobs or housing, are charged higher interest rates, are stopped more frequently by law enforcement, and face systematic disadvantages on online platforms (Bayer et al., 2016; Bertrand & Duflo, 2017; Bertrand & Mullainathan, 2004; Chetty et al., 2020; Doleac & Stein, 2013; Edelman et al., 2017; Fryer, 2019). Most of these disparities persist over generations, resulting in a lock-in effect of ethnic minority members, accumulating disadvantages, such as school quality, neighborhood conditions, health outcomes, and work experience over time (Alsan et al., 2019; Bayer & Charles, 2018; Bertrand & Duflo, 2017; Chetty et al., 2014). Such cumulative effects make it exceedingly difficult for ethnic minority members to escape entrenched inequalities and attain long-term improvements in life outcomes.

One might argue that lower productivity would economically justify lower earnings, that lower reliability would warrant higher interest rates or rents, or that a higher propensity for criminal behavior would rationalize more frequent stops and searches by law enforcement officers. However, a substantial body of research indicates that, once education quality, work experience, underwriting standards, or other background characteristics are controlled for, ethnic minorities do not exhibit systematically lower performance, greater default risk, or inherently higher criminal tendencies (Altonji & Blank, 1999; Donohue III & Levitt, 2001; Munnell et al., 1996; Neal & Johnson, 1996). On the contrary, the observed differences in outcomes are less due to actual deficits in skills, reliability, or others, rather than to biased beliefs, perceptions and stereotypes of the decision-makers (Bertrand & Duflo, 2017; Bertrand & Mullainathan, 2004; Bordalo et al., 2016; Coffman et al., 2023). In principle, this should allow for the rectification of biased beliefs, stereotypes, or others.

Today's social media and other online platforms – ranging from networking sites to job boards and rental sites – frequently serve as gatekeepers. While in general, platforms offer the potential for efficient matching, networking, and transparency, they can also allow biases to persist at scale through automated filtering, biased algorithms, or user-driven review systems (Aneja et al., 2025; Lambrecht & Tucker, 2019; Luca, 2016). Understanding how information from social media and, more general, online (matching) platforms shape, transmit, or curb discriminatory behavior remains a major frontier in research on inequality and digital markets.

This raises several important questions about the effects of using information from social media and online matching platforms on discrimination, as well as the mechanisms and persistence of *digital* discrimination. In which ways does additional (stereotypical) information about ethnic minority applicants ameliorate or reinforce biased beliefs and stereotypes on online matching platforms? How can unequal treatment be effectively reduced? And, in an era dominated by social media and online platforms, which new opportunities or challenges emerge concerning discriminatory practices, and how do these digital contexts shape decisions and behavior of both users and advertisers, landlords, employers, or others?

The first three chapters of this thesis (Chapters 2, 3, and 4) shed light on these questions by investigating the effects of ethnicity, gender, (stereotypical) social media information, as well as the salience of applications for vacant rooms in shared apartments and personal social network formation on discriminatory behavior. Large parts of economic research ignored the effects of (visual) minority stereotypes. This is despite the fact that psychological research suggests that providing accurate, personalized, or stereotype-disconfirming information can help reduce prejudice, in part by prompting decision makers to update their priors in line with new data rather than entrenched biases (Beaman et al., 2012; J. C. Becker & Swim, 2011; Bohren et al., 2019; Pettigrew & Tropp, 2006). Nonetheless, the extent and durability of such "information effects" remain contested, since biases can be deeply embedded and susceptible to motivated reasoning. Furthermore, there is limited causal evidence concerning how today's information channels – particularly those involving searches or "scouting" of social media users – shape both online and offline outcomes in housing markets or personal social network formation.

Previous literature concentrates in almost all cases on professional markets, such as job search or professional networking (Acquisti et al., 2015; Ajzenman et al., 2025; Evsyukova et al., 2025; Manant et al., 2019). In contrast, my research focuses on everyday economic and social exchanges in informal settings. While professional markets, such as the labor market, have important implications for economic growth and welfare, informal markets likewise affect job search and professional networking but extend more broadly to encompass diverse aspects of social life – from searching a room or an apartment, to forming friendships, relationships, and broader social affiliations – and may therefore play an even more critical role in tackling persistent inequalities.

Research on discrimination typically distinguishes between two main strands: taste-based (G. S. Becker, 1957) and statistical discrimination (Aigner & Cain, 1977; Arrow, 1973; Phelps, 1972). Taste-based discrimination arises when decision-makers hold prejudicial preferences against members of certain groups, independent of any productivity or quality-related trait. In contrast, statistical discrimination emerges when decision-makers resort to group-level heuristics or stereotypes to fill information gaps about an applicant's characteristics or future performance. Online platforms, whether social media or marketplace websites, are places where such discrimination likely plays out. On the one hand, the wealth of personal information on these platforms could reduce statistical discrimination by enabling more accurate assessments of applicants. On the other hand, these same cues could activate biases if they perpetuate stereotypes or confirm negative group-based expectations.

Building on these theoretical insights and drawing on large-scale randomized field experiments in searching for rooms or friends, the research in Chapters 2 to 4 explores whether and to what extent additional social media information about personal background, lifestyle, or personality can reduce, eliminate, or exacerbate discriminatory behavior. Collectively, they test the limits of technology-driven transparency: Can a well-curated social media profile counteract negative stereotypes associated with an individual's name or ethnic appearance? Or might new visibility cues (e.g., photos that confirm common minority stereotypes) amplify bias, making it harder for minority applicants to gain acceptance in social networks or to find housing?

The research in Chapters 2 to 4 aims to shed light on these questions by providing causal estimates of how behavior changes when additional (stereotypical) social media information is introduced or salience of an application is varied. This research contributes to the existing literature in several ways: Firstly, by investigating how ethnic minority status, gender, and (stereotype-free) social media information affect ethnic and gender discrimination in searching for a vacant room (Chapter 2), we demonstrate how an update of initial (stereotyped) beliefs reduces discriminatory behavior – an important advantage over previous studies in professional markets (Acquisti & Fong, 2020; Manant et al., 2019). Secondly, by investigating how minority status and (stereotypical) social media information signaling cultural identification and religious affiliation affect the formation of informal social networks and the search for a room (Chapter 3). This research is the first to incorporate (visual) stereotypes to causally identify the role of stereotypical social media information on market outcomes. In addition, previous literature on network formation concentrated on professional settings only (Ajzenman et al., 2025; Evsyukova et al., 2025). Thirdly, we investigate how minority status, (stereotypical) social media information, and the salience of an application affect ethnic discrimination (Chapter 4). Prior research (e.g., Acquisti & Fong, 2020; Bartoš et al., 2016; Bertrand & Mullainathan, 2004) lacks the ability to manipulate how résumés or applications are ordered and presented, making it difficult to cleanly identify the causal impact of an application's salience. Moreover, this research contributes to the growing body of field experiments exploring how digital platform features interact with discrimination (Aneja et al., 2025; Doleac & Stein, 2013; Edelman et al., 2017).

Thereby, this research investigates discrimination in the formation of informal social networks, the housing market, and, more generally, how minority cues, salience and stereotypes contribute to inequality in non-professional, informal domains. The rather informal market for shared housing[1] has several advantages to investigate the role of social media information on digital discrimination. Roommates are rather young, frequently use social media and other online platforms, and are willing to disclose personal information online (de la Llama et al., 2012), exploiting social media information does not violate anti-discrimination laws, and applications for vacant rooms often include social media profiles which makes it possible and realistic to integrate links to (fictitious) personal social media profiles directly into the applications – in contrast to previous literature (Acquisti & Fong, 2020; Manant et al., 2019).

In Chapter 2, we apply for about 3,700 vacant rooms in shared apartments randomly assigning ethnicity, gender, and whether the application contains a link to a corresponding social media profile that is apt to break with prevailing ethnic stereotypes. In Chapter 3, we use the very same profiles, but add visual stereotypes, i.e., images signaling common ethnic minority

[1] The majority of students in Germany live in shared apartments (Kroher et al., 2023).

stereotypes as held by the ethnic majority, to some of the profiles and apply to roughly 3,100 vacant rooms. In addition, we use the manipulated profiles to send about 1,000 friend requests to real users to analyze how ethnicity and visual stereotypes affect personal social network formation. Subsequently, in Chapter 4, we add another treatment arm to the experimental design from Chapter 3, and randomly select room applications to be send via a subscription-based premium service that highlights and ranks the application near the top of the roommate's/advertiser's inbox. We send roughly 4,300 applications to vacant rooms to investigate whether enhancing the salience of an application can mitigate or exacerbate ethnic discrimination on digital platforms.

The results of Chapters 2 to 4 for about 6,900 applications for vacant rooms indicate an average callback rate of 35 percent for ethnic majority applicants versus 23 percent for applicants with a minority name – indicating an ethnic gap of 52 percent. Thus, minority members have to send 52 percent more applications on average to receive one callback compared to majority members. Figure 1.1 indicates that this gap increases over time, as in Chapter 2 it is 31 percent (the experiment was conducted in 2021), in Chapter 3 it is 60 percent (2023), and in Chapter 4, it is 71 percent (2024).[2] These differences are statistically significant at the one percent level.

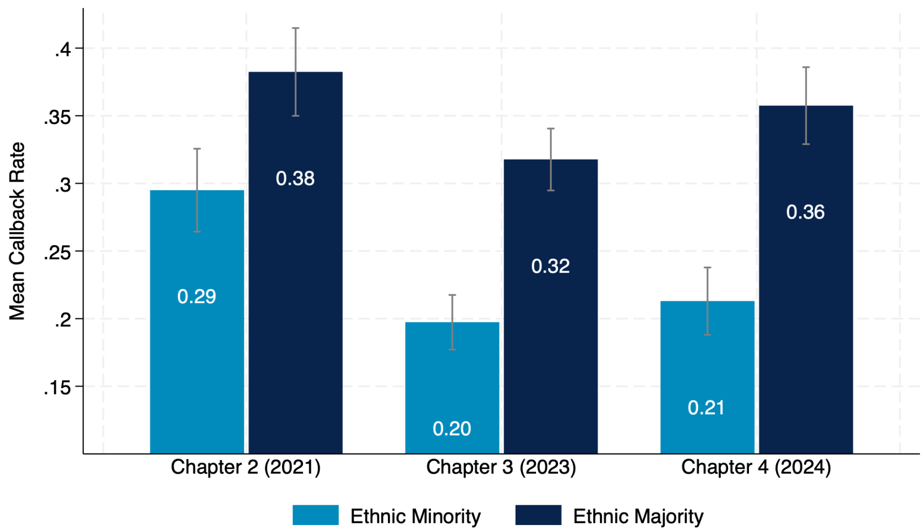


Figure 1.1: Mean Callback Rates per Chapter

Regarding different social media information treatments, the results from Chapters 2 to 4 indicate that applications attaching a social media profile *without* minority stereotypes exhibit the lowest ethnic gap (37 percent), compared to 65 percent for applications attaching a social media profile *with* minority stereotypes, and 64 percent for applications without an additional social media profile. Importantly, the results of Chapter 2 indicate that social media profiles free

[2]The mean callback rates are computed using data from Chapters 2, 3, and 4 for applications for vacant rooms, excluding female names (from Chapter 2) and premium applications with higher salience (from Chapter 4), to ensure comparability. In addition, the data of Chapter 2 is not restricted to ads that do not impose gender restrictions (see Section 2.4.1), again to ensure comparability. Excluding treatments with stereotypical social media profiles (Chapter 2 only features stereotype-free profiles) does not considerably change the results displayed in Figure 1.1.

of minority stereotypes, i.e., “breaking” biased beliefs or stereotypes, lead to an elimination of the ethnic gap.

In terms of personal social network formation, the findings in Chapter 3 reveal similar patterns for friend requests as for roommate searches. Overall, minority profiles receive 46 percent fewer friend acceptances. Displaying stereotypical photos on one’s profile reduces acceptance rates significantly: Minority profiles featuring stereotypical images experience a 69 percent lower acceptance rate than majority profiles. Again, discrimination is lowest (31 percent gap) for stereotype-free profiles. All differences are statistically significant at the one percent level.

An experimental manipulation of the salience of an application, operationalized by an application’s position in the advertiser’s or roommate’s inbox (see Chapter 4), shows that while greater salience slightly improves callback rates for both minority and majority applicants, it does not affect the persistent, significant ethnic gap. The ethnic penalty also remains unaffected by social media contents, i.e., whether minority stereotypes are shown on a profile or not. Moving one position lower in the inbox reduces an applicant’s callback probability by 0.27 percentage points, an effect that becomes substantial when spanning multiple positions and disproportionately affects minority applicants – particularly in inbox’ positions where applications are still likely to be read rather than dismissed.

Overall, the findings from Chapters 2 to 4 highlight the pervasiveness of ethnic discrimination in informal online contexts, such as searching for friends or roommates. While challenging existing ethnic stereotypes on social media can eliminate or at least narrow the ethnic gap, providing no social media information can, in some cases, prove as detrimental as providing a stereotypical profile. Furthermore, the data reveal that social media profiles are actively accessed and affect decision-making, with ethnicity significantly affecting all engagement metrics. Notably, however, attention-based mechanisms and enhanced application salience fail to mitigate ethnic discrimination, indicating that increasing visibility alone does not overcome entrenched biases.

Personality and Social Media While Chapters 2 to 4 focus on the effects of ethnicity and stereotypes in searching friends and roommates online, Chapter 5 investigates the effects of visual personality cues on real-world outcomes, i.e., friend acceptance and callback rates using two randomized field experiments.

Social media platforms provide extensive insights into users’ personalities and play a central role in predicting individual outcomes across various contexts (Back et al., 2010; Barrick & Mount, 1991; Judge et al., 1999; Kluemper et al., 2012; Rothmann & Coetzer, 2003). Numerous studies have investigated how self-, friend-, and system-generated information on social media profiles affect impression formation (Utz, 2010; Walther et al., 2009; Walther et al., 2008). However, less is known about how personality cues derived from social media profiles affect real-world outcomes. This gap is especially pertinent because social media not only serves to maintain existing relationships but also facilitates meeting new people or screening potential roommates, colleagues, or employees. Much of the existing research relies on hypothetical survey responses or small-scale hypothetical laboratory tasks (Domahidi et al., 2022; Utz, 2010; Walther et al., 2009; Walther et al., 2008), which significantly limits the external validity.

Chapter 5 addresses this gap by conducting two field experiments using genuine online interactions, ensuring that decision-makers (social media users and roommates) make real choices with tangible consequences. The design allows us to causally identify the effect of fictitious social media profiles that systematically manipulate (1) the agreeableness/emotional stability of the fictitious profile owner or (2) her conscientiousness – both traits play an important role in predicting individual outcomes in organizational and other contexts (Barrick & Mount, 1991; R. Fang et al., 2015; Judge et al., 1999; Rothmann & Coetzer, 2003) – by sending friend requests to real Instagram users and applications to vacant rooms using the fictitious social media profiles. The resulting evidence highlights the causal impact of these personality traits on real-world outcomes.

The results of approximately 2,800 applications for vacant rooms and 1,000 friend requests on Instagram reveal that profiles signaling high agreeableness/emotional stability elicit significantly higher acceptance and callback rates than those signaling low agreeableness/emotional stability. Differences for high versus low conscientiousness are not statistically significant. The findings underscore the important role of agreeableness/emotional stability for interpersonal interactions, regardless of the setting.

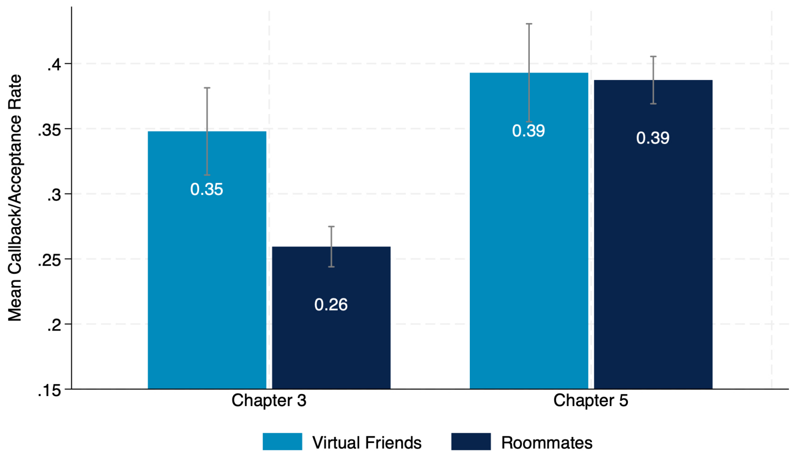


Figure 1.2: Mean Callback & Acceptance Rates per Chapter

The effects are very similar across both contexts, searching for friends or roommates (see Figure 1.2)[3] – unlike the results in Chapter 3, where the difference between acceptance rates for virtual friends and callback rates for roommates is large and significant.

ESG Metrics and Executive Compensation While Chapters 2 to 5 center on social media signals at the individual level, Chapter 6 turns to a different research question: how organizations incorporate non-traditional performance metrics into executive compensation and the implications regarding executive pay risk – clarifying whether such novel performance metrics genuinely transform corporate behavior or merely function as "greenwashing" or "window dressing".

[3]The mean callback and acceptance rates are computed using data from Chapters 3 and 5 for friend requests and applications for vacant rooms. The data is pooled, i.e., combines the ethnic minority and majority samples from Chapter 3 and high vs. low levels of conscientiousness and agreeableness/emotional stability from Chapter 5. The results presented here remain largely unchanged when ethnicity and personality traits are analyzed separately.

Over the last decade, there has been growing pressure on firms to account for broader societal and environmental impacts (Hazarika et al., 2022; PricewaterhouseCoopers, 2021; Willis Towers Watson, 2023). Environmental, social, and governance (ESG) metrics have thus become increasingly prominent in executive pay structures, ostensibly to align leadership decisions with stakeholders' diverse interests – taking shareholders, employees, communities, and the environment into account (Cohen, Kadach, Ormazabal, & Reichelstein, 2023; Eccles et al., 2014; Ikram et al., 2023). This expansion of performance measures beyond purely financial targets may be viewed as similar to the social media phenomenon from previous chapters: in both spheres, new forms of information are introduced with the promise of driving more equitable, informed, and/or responsible decision-making.

Despite the enthusiastic uptake of ESG metrics, concerns about greenwashing or window dressing persist (Kolk & Perego, 2014; Yu et al., 2020). Critics worry that ESG metrics might be too vague, too discretionary, or too lightly weighted in compensation formulas to have a meaningful impact on executive behavior (Cohen, Kadach, Ormazabal, & Reichelstein, 2023; Homroy et al., 2023; Maas, 2018). If such metrics are included primarily to signal social responsibility rather than to provide binding incentives, then they may not address the underlying issues that they supposedly aim to improve.

Chapter 6 tackles these issues by analyzing a detailed hand-collected panel dataset of European executives to observe how ESG metrics are integrated into compensation contracts. The previous literature predominantly analyzed commercially available data sources (e.g., Bebchuk & Tallarita, 2022; Cohen, Kadach, Ormazabal, & Reichelstein, 2023). However these data exhibit significant shortcomings; e.g., many only include a single dummy variable that indicates whether there are any ESG related metrics in the compensation contract. This research, on the other hand, collects data on each and every metric and its weights, classify them into financial, non-ESG, and ESG-targets, assesses target achievement, and merges this data with realized pay of executives.

Therewith, Chapter 6 can robustly identify whether ESG measures are truly binding – i.e., carry a substantial payout weight and thus meaningfully affect overall pay risk and can hence be expected to affect executive behavior and corporate strategy – or whether they remain largely symbolic. The analysis includes (1) the prevalence and characteristics of ESG metrics by investigating which firms adopt ESG-based incentives, how these metrics are structured, and what their relative importance is compared to traditional performance measures (e.g., share price, earnings); (2) the potential for greenwashing by assessing whether the prevalence of discretionary or weak ESG metrics is higher in certain sectors (e.g., financial firms, large companies, high-visibility executives) that may be under public pressure to appear socially responsible without making significant changes; and (3) the role of such metrics in industries or sectors with high environmental footprint by identifying whether such firms are more likely to use robust, high-weighted ESG metrics – suggesting greater alignment with genuine sustainability objectives.

The findings suggest that while ESG metrics are becoming more common, many are designed in a way that imposes minimal additional pay risk on executives, raising questions about their ability to bring about real behavioral change.

These findings parallel the insights from previous chapters about *non-traditional* information: introducing new forms of information, such as social media signals on (online) platforms or non-financial performance metrics in incentive contracts, do not necessarily lead to a behavioral change or transformation that foster more equitable or socially responsible outcomes.

On both, online platforms and in boardrooms, market frictions and decision biases can hamper effectiveness. On the one hand, (digital) discrimination often emerges from stereotypes, implicit biases, or overreliance on group-level heuristics. On the other hand, a form of "greenwashing" may arise in organizational settings if sustainability aspirations fail to align with executive's pay risk and therefore, fail to induce behavioral changes. In each case, decision-makers – whether individuals in informal settings, such as social network users, potential roommates, or members of corporate boards designing incentive contracts – may inadvertently perpetuate inequalities or inefficiencies if these information signals are under weighted or strategically manipulated. Thus, how information and signals are received, used, and processed is critical in determining its effectiveness.

Despite these challenges, this thesis' findings also highlight opportunities for (partial) mitigation. Chapters 2 to 4 show that when social media cues are carefully designed to break prevailing stereotypes, they can reduce discriminatory behavior, offering tangible evidence that biased outcomes can be moderated, if not entirely eliminated. Furthermore, Chapter 4 demonstrates that not only the nature of (non-traditional) information and signals matters, but also their salience. However, paying a premium for increased information salience fails to reduce the persistent ethnic gap. Similarly, Chapter 5 shows that personality cues significantly affect both online friend acceptance and roommate selection to an equal degree. Chapter 6 reveals that in sectors with large environmental footprints, ESG metrics often carry more meaningful weights, indicating that external scrutiny, regulatory pressures, and reputational concerns encourage stronger incentive structures. Furthermore, when metrics, or signals, are implemented without robust incentives or genuine accountability, their transformative potential remains limited, mirroring the dynamics observed in previous chapters' field experiments.

Altogether, the insights from Chapters 2 to 6 present a comprehensive examination of how individuals and organizations respond to *non-traditional* information. By conducting large-scale field experiments, this thesis explores mechanisms through which biases and behaviors evolve – or persist – in the face of novel signals. The results emphasize the importance of careful design, credible signaling, and accountability in harnessing non-traditional information. It invites continued exploration into how to design interventions and incentives that genuinely foster inclusion, equity, and responsible stewardship in both digital and corporate settings.

Chapter 2

Can Social Media Information Reduce Discrimination?[1]

2.1 Introduction

Social media adoption has never been more pervasive, with approximately more than four billion registered users of Social Network Sites (SNSs) in 2021 – an all-time high (Datareportal, 2021). Among others, SNSs provide a channel for disclosing information that is otherwise inaccessible to potential contractors (Acquisti & Fong, 2020; Kluemper et al., 2012), and thus could alleviate problems associated with asymmetric information about qualities of the profile owner that are deemed relevant for contractual exchange in a two-sided market (Akerlof, 1970). For instance, in the market for shared housing, our field of application, the potential roommates will want to know whether an applicant for a vacant room has a desired quality, such as a certain level of sociability or orderliness (Clark et al., 2018; Diehl et al., 2013). In the absence of any other source of information, the uninformed side of the market, i.e., the incumbent residents, will form conditional probabilistic beliefs based on (known or perceived) statistical correlations between applicants' observable characteristics (e.g., ethnicity) and the desired qualities (Arrow, 1998; Lang & Kahn-Lang Spitzer, 2020; Phelps, 1972).

While forming expectations in such a way may be perfectly rational (Arrow, 1973; Phelps, 1972), it inevitably leads to (statistical) discrimination where individuals who are identified as members of a certain group (e.g., in terms of their ethnicity) have lower chances of market participation because, overall, members of that group are less likely to share a certain quality (Aigner & Cain, 1977; Allport, 1954; Lee et al., 2015). In this context, SNSs can provide additional information that leads the uninformed side of the market to update its probabilistic beliefs, thereby alleviating problems of asymmetric information and thus reducing (statistical) discrimination. This is where our study comes in.

[1] Chapter 2 is based on the paper "#InviteMe: Can social media information reduce discrimination? Evidence from a field experiment" by Raphael Moritz (University of Tübingen), Christian Manger (University of Tübingen), and Kerstin Pull (University of Tübingen). This paper has been published in the Journal of Economic Behavior & Organization *213*(2023): 373-393. https://doi.org/10.1016/j.jebo.2023.07.032. The experiment was approved by the Ethics Committee of the Faculty of Economics and Social Sciences at the University of Tübingen in July 2019 (IRB approval A2.5.4-083_aa). The experiment has been pre-registered in the AEA RCT Registry (#7627): https://doi.org/10.1257/rct.7627-2.0.

Unlike Manant et al. (2019) and Acquisti and Fong (2020), who find that SNSs can provide an additional *source* of discrimination, we show that social media information can also be used to reduce (statistical) discrimination. Our finding differs from Manant et al. (2019) and Acquisti and Fong (2020) because, in our study, the fictitious SNS profiles were created to induce an update of beliefs by deliberately breaking with ethnic stereotypes that the uninformed side of the market may hold.

In our paper, we conduct a randomized field experiment to study whether and to what extent SNSs can reduce (statistical) discrimination. For this purpose, we create eight fictitious SNS profiles (male and female) on the photo-sharing platform Instagram. Except for the names of the profile owners, the four male and four female profiles are identical. The only difference between the profiles of one gender is the name assigned to them: two profiles of each gender are assigned a Turkish-sounding name, while the other two are assigned a German-sounding name. Each application is randomly assigned one of the eight names and whether it contains a link to the corresponding Instagram profile. To make the social media information easily accessible to the potential roommates or advertisers, the link to the SNS profile is included directly in the application text, thus reducing the search costs of the uninformed side of the two-sided market to virtually zero.

In contrast to the studies by Manant et al. (2019) and Acquisti and Fong (2020), the alleged ethnicity of an applicant is already indicated by the applicant's name in the application text and not only disclosed in the SNS profile. Thus, our experiment is the first to investigate whether and to what extent social media information can *reduce* (statistical) discrimination. Unlike previous studies, our experiment includes both female and male names, which allows us to analyze whether the effects of including an SNS profile that breaks with prevailing ethnic stereotypes, i.e., by not showing any stereotypes that the ethnic majority holds about the ethnic minority, differ between the two genders. Finally, we are the first to be able to analyze the profile statistics of our fictitious applicants in order to verify whether and to what extent our fictitious SNS profiles were actually visited during the course of our experiment.

After creating the SNS profiles, we apply to 3,676 vacant room ads on the shared housing market in Germany where this form of housing is particularly popular among students (Hartmann & Knappe, 2019) with about 30% of students living in shared apartments (Middendorff et al., 2017). In the United States, more than 10% of all households are shared (United States Census Bureau, 2020). Studying the role of social media information in this market has several advantages: First, participants in this market are rather young, frequent users of SNSs, and willing to disclose personal information on SNSs (de la Llama et al., 2012). Second, in contrast to the labor market, the shared housing market is rather informal, so that information from (non-professional) SNSs can be more easily exploited, without, for example, violating anti-discrimination laws. Third, it is quite common in the shared housing market to include a reference to a non-professional SNS profile in the application text. Some advertisers even explicitly ask applicants to include it in their application – which would be rather unusual in more formal, professional markets.

In our analysis, we control for a large set of variables, including room and apartment characteristics, as well as variables on district level, such as the extent to which a district was affected by the Covid-19 pandemic at a given point in time.

Our results show that social media information can in fact reduce ethnic discrimination: Absent any SNS profile, the average callback rate for applicants with a German-sounding name is 41.96%, while it is only 31.84% for applicants with a Turkish-sounding name. When an SNS profile that is apt to break ethnic stereotypes about ethnic minority applicants is included, the callback rate is 43.07% for applicants with a German-sounding name (+1.11 percentage points) and 41.34% for applicants with a Turkish-sounding name (+9.5 percentage points). Thus, ethnic minority applicants substantially increase the likelihood of a callback compared to an ethnic majority applicant by including an SNS profile.

This effect is strongest for minority male applicants, who – absent an SNS profile – have a callback rate of only 25.22% and whose callback rate increases to 37.08% (+11.9 percentage points) when an SNS profile is included. The resulting callback rate is no longer statistically significantly different from the callback rate of the majority male applicants. That is, members of the most disadvantaged group (males with a Turkish-sounding name) benefit the most from including an SNS profile that is apt to break ethnic stereotypes in their application.

Furthermore, our data on SNS profile visits reveal that the profiles of our fictitious applicants are in fact visited and screened by potential roommates, so that the profiles can be used to update roommates' conditional probabilistic beliefs about whether the corresponding applicant shares a desired, but otherwise hidden quality.

The remainder of this chapter is organized as follows: Section 2.2 presents our theoretical framework and reviews the relevant literature. Section 2.3 describes the experimental setup, and Section 2.4 presents our results. Section 2.5 concludes.

2.2 Theoretical Framework and Literature

Theoretically, our study is embedded in a situation characterized by pre-contractual information asymmetry. More specifically, we refer to a hidden characteristics setting in a two-sided market where one side is not informed about relevant qualities of what the other side has to offer (Akerlof, 1970), and where the uninformed side thus has to form expectations about these hidden qualities (Arrow, 1973; Arrow, 1998). In a labor market setting, the uninformed side of the market typically refers to the employer who seeks to fill an open job position and who is not informed about relevant qualities of the applicants (Cornell & Welch, 1996). In our setting, the shared housing market, the uninformed side of the market refers to the incumbent residents of the apartment who are searching for a new roommate and who are likewise uninformed about relevant qualities of the applicants.

Due to time (and cost) constraints, the uninformed party in a labor market or in a shared housing setting will invite only a preselected number of applicants to an interview or apartment viewing, respectively. Whether a particular applicant is invited or not depends on the probabilistic beliefs of the uninformed side of the market about whether the applicant has the desired qualities (Cornell & Welch, 1996; Neumark, 2018). Lacking any other information, the uninformed side of the market will base its expectations on observable characteristics (Aigner

& Cain, 1977). In his seminal paper, Spence (1973) distinguishes between observable characteristics that are given (so-called “indices”, e.g., an applicant’s ethnicity) and those that are a matter of choice (“signals”, e.g., the number of years an applicant has attended school (Spence, 1973) or whether an appropriately designed SNS profile is attached to the application).

For a signal to be credible and hence apt to induce a separating equilibrium, signaling must be costly, and applicants with the desired quality must have lower costs of sending the signal than those who do not have the desired quality (Spence, 1973, 2002). That is, in a labor market setting, applicants with the desired quality need to have lower costs of investing in, e.g., education, and in our shared housing setting, applicants with the desired quality need to have lower costs of investing in an SNS profile that breaks with prevailing ethnic stereotypes (Pager & Karafin, 2009; Roth et al., 2016; Spence, 1973). While adding an SNS profile to the application is essentially free, adding a profile that is apt to break prevailing ethnic stereotypes may well involve costs, and these costs (e.g., in terms of having to be highly selective when uploading pictures) will typically be higher for those applicants who do not have the desired qualities.

Absent any such signals, the uninformed side of the market can rely solely on indices (e.g., an applicant’s ethnicity) when forming expectations about whether a particular applicant has the desired qualities (Spence, 1973), and it will do so based on (known or perceived) statistical correlations between, e.g., an applicant’s ethnicity and the desired qualities (Arrow, 1973; Phelps, 1972). While forming expectations in such a way may be economically rational, it inevitably results in statistical discrimination. Unlike taste-based discrimination, where discriminating agents have – unsubstantiated – reservations about members of a certain group in terms of, e.g. their ethnicity (G. S. Becker, 1957), statistical discrimination arises as a result of imperfect information (Aigner & Cain, 1977; Arrow, 1973; Phelps, 1972), where members of a certain group have lower chances of market participation because, they are statistically (perceived to be) more likely to share a certain quality (Aigner & Cain, 1977; Allport, 1954; Lee et al., 2015). In this setting, additional information can lead the uninformed side of the market to update its beliefs, thus reducing (statistical) discrimination.

Although discrimination on the basis of ethnicity, for example, is prohibited in many countries, there is an abundance of field experiments that provide ample evidence of ethnic discrimination in many different settings around the world (Bertrand & Duflo, 2017). In correspondence studies, fictitious applications that systematically vary selected characteristics of applicants are sent out to employers, landlords, and others, to find out whether there is discrimination as measured by the difference in callback rates for different groups of applicants (Harrison & List, 2004). As it is, an impressive number of these correspondence studies find ethnic discrimination in labor, (shared) housing, and other markets (Auspurg et al., 2019; Bertrand & Mullainathan, 2004; Carlsson & Eriksson, 2015; Carlsson & Rooth, 2012; Carpusor & Loges, 2006; Cui et al., 2020; Lippens et al., 2023; Nunley et al., 2011).

That a potentially substantial part of the observed ethnic discrimination is statistical rather than taste-based, is highlighted by several studies that analyze the effect of additional information provided on the extent of discrimination (Baert et al., 2017; Bosch et al., 2010; Kaas & Manger, 2012; Moritz & Manger, 2022). For instance, Kaas and Manger (2012) find that adding a favorable reference letter to a vacant job application eliminates the otherwise substantial discrimination against job applicants with an ethnic minority name as opposed to an ethnic majority name. Similarly, BAERT et al. (2017) find that adding information about applicants' work experience reduces discrimination against ethnic minorities. In the housing market, providing additional information has also been shown to reduce discrimination against ethnic minorities (BOSCH et al., 2010; MORITZ & MANGER, 2022). For instance, MORITZ and MANGER (2022) conduct a correspondence study in which they experimentally vary whether a brief (and favorable) self-description is added to the application of a (fictitious) applicant who is either a member of the ethnic majority (i.e., German) or the ethnic minority (i.e., Turkish). MORITZ and MANGER (2022) find that adding the short self-description to the application of a male minority applicant (who otherwise suffers from the lowest callback rates) significantly increases his chances of being called back, suggesting that a significant portion of the discrimination observed is statistical rather than taste-based.

To date, however, there is no empirical evidence on the potential role of *social media information* in reducing statistical discrimination. This is despite the fact that non-professional SNS profiles are an increasingly important source of information (BROWN & VAUGHN, 2011) with a significant proportion of employers using SNS to scout potential candidates (BECTON et al., 2019; ROSEN et al., 2018) or to assess the employability of candidates (ROTH et al., 2016).

It is only recently, that two studies have explored the role of information provided on SNS profiles in the context of ethnic discrimination (ACQUISTI & FONG, 2020; MANANT et al., 2019), but both of these focus on social media information as an additional *source* of discrimination, and their designs do not allow to infer to what extent social media information might *reduce* (statistical) discrimination by *differentially affecting* the callback rates of ethnic minority applicants as opposed to ethnic majority applicants.

ACQUISTI and FONG (2020) send out 4,000 job applications in the U.S. using distinctive male names for which they create non-professional SNS profiles. On social media only, they specify personal characteristics related to religious beliefs (Christian vs. Muslim) and sexual orientation (straight vs. gay). Overall, they find no significant differences in callback rates for any of the characteristics. However, in areas with a high share of Republican voters, the fictitious Christian applicant receives significantly more callbacks than the fictitious Muslim applicant. As the information about religious beliefs is only provided on the SNS profiles, and because ACQUISTI and FONG (2020) know from supplementary data (e.g., an additionally created professional SNS profile) that employers in fact searched for the names of the fictitious applicants online, the authors conclude that the SNS profiles lead to discrimination.

MANANT et al. (2019) create two Facebook profiles and assign these two unique first and last name combinations of a fictitious male. One of the two Facebook profiles indicates an Arab origin, the other a French origin. This information is not included elsewhere in the (otherwise identical) application material that the authors send to approximately 800 job openings in France. MANANT et al. (2019) find a 37% difference in callback rates favoring the candidate with the alleged French origin over the one with the alleged Arabic origin. An exogenous change in the layout of the profiles results in a lower salience of the information about an applicant's ethnic origin, which eliminates the difference in callback rates. Based on this latter finding and the fact that the information on ethnic origin is only provided on Facebook and nowhere else, Manant et al. (2019) conclude that SNS profiles are indeed assessed (albeit not screened in great detail) by employers and that social media information leads to discrimination.

Given the design of the studies by Acquisti and Fong (2020) and Manant et al. (2019), a potential differential effect of social media information, depending on whether the SNS profile is added to the application of a member of the ethnic minority group vs. the ethnic majority group, cannot be assessed. Similarly, a potential variation by gender, for example tested by Moritz and Manger (2022), cannot be examined. Furthermore, the two existing studies are silent on whether the induced discrimination is statistical or taste-based. Finally, both of the existing studies can only provide rather indirect evidence on social media information being assessed by the uninformed side of the market. With our study, we seek to address all of these issues.

Furthermore, our study differs substantially from the work of Moritz and Manger (2022) who also analyze the role of additional information on ethnic discrimination in the shared housing market: First, instead of text-based information, we study the role of *visual* information posted online on a fictitious social media profile of the applicant. With the emergence of newer SNSs such as *Tiktok*, the relevance of text-based information in social media has steadily decreased. Second, and equally important, we focus on information that is publicly available online and thus more likely to serve as a signal in the sense of Spence (1973): Describing oneself in a favorable way in a letter to the advertiser, which does not correspond to the "true self" of the applicant, is essentially costless and *cheap talk*. To the contrary, creating an inauthentic social media profile will be costly, (a) because applicants would have to always think twice about which photo (not) to upload to their profile, and (b) because applicants will not want their friends, family, or colleagues to see a profile that does not reflect their true self and identity. Correspondingly, previous research suggests that social media profiles are more likely to reflect the true personality rather than an idealized version of the profile owner (Back et al., 2010; Kluemper et al., 2012).

2.3 Experimental Design

We study discrimination in selection using a between-subjects, non-matched correspondence test design to apply for vacant room ads. Each advertiser is randomly assigned to receive one application from a fictitious applicant with either a Turkish- or a German-sounding name (male or female), with or without an additional link to an SNS profile.

2.3.1 Fictitious Applicants

Germany, like the United States, is considered an immigration country (Thränhardt, 1995). In 2021, 22.3 million people in Germany have a migration background, representing about 27% of the total population (Federal Statistical Office, 2022b). About 12% (2.7 million) of these are people with a Turkish migration background. This makes Turkish people the largest foreign population group in Germany (Federal Statistical Office, 2022c).

We build on the selection of the most popular Turkish and German female and male names of the respective birth cohorts by Moritz and Manger (2022). To control for differences in names, we select the two most "average" names in terms of callbacks from their study. This results in a total of eight names, such as "Zeynep Yildirim" and "Muhammed Kaya" for Turkish applicants, and "Lisa Müller" and "Tobias Weber" for German applicants.[2] For each of these names, there are at least 50 SNS profiles, thus making it unlikely that room advertisers who receive an application without a link to an SNS profile will screen them in search of our fictitious applicants.

Each advertiser of a vacant room is confronted with the name of the fictitious applicant twice in the application text.[3] In addition, at the beginning of the message, the online platform for shared apartments displays the applicant's name above the text, including the e-mail address, which is also a combination of the applicant's first name and surname. In total, the name is visible to the advertiser at least four times. For those applications that additionally contain a link to an SNS profile, the applicant's username and the profile link also consist of the applicant's name. Hence, it is almost impossible to overlook an applicant's name indicating ethnicity and gender.

All fictitious applicants are 24-year-old students pursuing a master's degree in business administration. The applicants' hobbies are consistent with the most frequently mentioned hobbies in shared housing application texts, i.e., meeting friends, jogging, and watching TV series, as well as with content frequently posted on SNS profiles, i.e., meeting friends and traveling (Hu et al., 2014). We also include an (exogenous) reason why our fictitious applicants are searching for a new room, referring to the fact that after the peak of the Covid-19 pandemic, lectures are being held on site again and the need to move. Applicants have previous experience living in a shared apartment and are non-smokers. For the randomly assigned SNS condition, we add a postscript that provides a link to an SNS profile, thus reducing search costs for the room advertisers to virtually zero.

2.3.2 Social Network Site Profiles

The fictitious SNS profiles are designed to be able to lead to an update of advertisers' potentially stereotypical beliefs against the dominant ethnic minority based on the applicants' alleged ethnicity as indicated by the applicant's names. Accordingly, we create SNS profiles that do *not* foster any of the prevailing stereotypes that the ethnic majority holds about the ethnic minority, and thus are apt to break prevailing stereotypes. More specifically, the fictitious profiles do not contain any information about a potential religious affiliation or religiosity, they do not show family orientation, do not allude to Turkish traditions, or national pride regarding Turkey – which represent the most common stereotypes that Germans hold about Turkish people (Ossenberg, 2019).[4]

[2] In addition, we verify that the names were correctly assigned to the intended ethnicity and gender by conducting an additional randomized online experiment ($n = 1,725$) in response to an anonymous reviewer's comment. Female (male) names were correctly classified as female (male) in the vast majority of cases (87.40 – 99.13 percent, depending on the specific name). Similarly, Turkish-sounding (German-sounding) names were also correctly assigned to the respective ethnicity in the vast majority of cases (93.09 – 99.66 percent, depending on the specific name).

[3] See Chapter A.1 (p. 150) for a translated version of the application text.

[4] To verify that our experimental target group (i.e., advertisers of rooms in shared apartments) holds similar stereotypes, we used our additionally conducted online survey to elicit stereotypes among students and found very

We create the fictitious SNS profiles on Instagram – a popular photo-sharing platform used mainly by young people and students (de la Llama et al., 2012; Frison & Eggermont, 2017). Registered Instagram users can upload photos and videos, add descriptions, and share them with other users (Hu et al., 2014).

Based on data collected from existing Instagram profiles, we design two types of SNS profiles (female and male) that would be considered to be rather "average" for a 24-year-old native German student. We then register eight profiles on Instagram, using the four female names and four male names. According to Hu et al. (2014), the most frequently uploaded photos on Instagram show the person who owns the profile, alone or with friends. Therefore, we recruit five students to provide us with personal social media images. In a randomized online pilot survey among students, we ask our survey participants to rate a candidate's personality (Rammstedt et al., 2020), perceived ethnicity, and whether they would share an apartment with the candidate based on a random selection of that candidate's photos. Based on the results of this pilot study, we select one female and one male candidate for our study who are very similar in all dimensions (regarding personality and whether they would qualify as roommates), *and* who could be perceived as being of both German and Turkish origin. The latter is very important since the four female and the four male Instagram profiles are identical except for the names of the profile owners.

Besides photos showing the profile owner alone ("selfies") or with friends, we also upload "filler" photos to the Instagram profiles, such as food or activity-related images, such as travel, nature, or indoor activities, that do not contain individually recognizable faces. These "filler" photos are kept constant across conditions. The final selection results in a total of 31 posts. We complete the Instagram profiles by adding biographical information such as age, student status, and the place of residence at the top of the profile page.[5]

The first image is uploaded in July 2019 and the last in May 2021, before the start of the experiment. We make sure that each photo is uploaded to all accounts on the exact same day. We add short, matching captions including hashtags to each photo, written by student assistants who also use SNSs themselves.[6] Some additional photos and short videos are added to the Instagram profiles as *Story* items. These stories show activities close to the place of residence such as travel, nature, and food photos in line with other frequently posted photos (Hu et al., 2014).

Further, we create an additional set of eight profiles (four female, four male) representing business administration students similar to our fictitious applicants and connect them to our candidates' profiles. Since these profiles are not publicly accessible, we simply ask some students to subscribe and post some photos in order to make the profiles somewhat realistic, although detailed information cannot be accessed by other users. We then mention these additional profiles in some of the photos of our fictitious profile owners where "friends" are shown.

similar results.

[5] See Figures A.1 (p. 151) and A.2 (p. 151) for screenshots of the profiles (female and male).

[6] Hashtags are words preceded by a # and typically describe the content of photos to enhance social connectivity (Hu et al., 2014). Hashtags can be used to find similar content that uses the same hashtag. We mainly use relatively short descriptions, like "Finally on vacation! #recreation #sun #sea" for a travel photo or "Homemade sushi #delicious #enjoy #healthy #tasty #studentlife" for a food picture.

Throughout the process of designing the profiles, we conduct multiple randomized online surveys to assess the authenticity of the fictitious Instagram profiles and ask for comments on their validity, recruiting $n = 306$ student participants. 87% of participants state that the profiles are extremely or very realistic. We also ask participants to subscribe to some of the fictitious Instagram profiles and to like a random selection of images to make the profiles as realistic as possible. We also have our fictitious profile owners subscribe to the new subscribers. Over the course of the experiment, each of our fictitious profiles averages about 109 subscribers, 172 subscriptions, and 937 image likes.

Finally, we enable the *professional profile* setting. This setting is not displayed publicly on the profile and provides access to extensive profile statistics on visits, impressions, and many other profile and post level variables. Professional profiles are automatically set to public so that most information (name, profile picture, biographical information, and posts) are accessible to any visitor.[7] This allows anyone to screen the profile at basically no cost.

2.3.3 Conducting the Experiment

The experiment was conducted between July and December 2021 in the fifteen largest student cities in Germany[8] on a two-sided platform, the largest website for shared apartment ads in Germany, "wg-gesucht.de"[9] – which is similar to, for example, Airbnb, but for long-term rentals. The website provides detailed information about room, apartment, and incumbent roommates' characteristics, such as rental and utility costs, size, gender and number of incumbent roommates, geographic location, type of cohabitation, etc.

We apply to all ads that are online for no more than two hours, due to the large number of applications that advertisers in large college towns typically receive within the first few hours of posting an ad. This process creates a fairly balanced sample. We do not apply for rooms that were offered by rental companies or real estate agencies, as selection decisions may differ from "regular" shared apartments, where non-professionals decide whether to call back an applicant based on the applicant's perceived quality/fit.

In addition, we exclude rooms that are only available for an interim period of less than 120 days, which is equivalent to a semester without exam periods. Furthermore, we do not apply for rooms that have an age preference that does not match our fictitious applicant, or that have other requirements that do not match our fictitious applicant.[10] Further, we adhere to any specification regarding the gender of the roommate being sought, i.e., to "female-only" ads, we send an application of a fictitious female applicant, and to "male-only" ads, we send an

[7] Additional information, such as descriptions of posts, stories, or followers, can only be viewed by registered users.

[8] More specifically, the experiment was conducted in Berlin, Hamburg, München, Köln, Frankfurt am Main, Stuttgart, Düsseldorf, Leipzig, Münster, Bochum, Aachen, Darmstadt, Giessen, Dresden, and Göttingen. In each of these cities, there is a university that offers a master's program in business administration or a related field, such as international business.

[9] As of September 2021, wg-gesucht.de had 14.3 million visitors with 106.2 million page impressions (see https://www.wg-gesucht.de/ic/online-werbung.html [Retrieved: July 29, 2022]).

[10] Typical requirements are that only people who are already working should apply, or that the application text should mention the applicant's favorite movie or food to verify that the applicant read the ad carefully.

application of a fictitious male applicant. If the gender of the roommate that is being looked for is not specified in the ad, one out of eight applications will be randomly assigned.

Each advertiser receives exactly one application. The application texts are identical across conditions, differing only in the applicant's name, which indicates ethnicity and gender, and in whether or not the application additionally contains a link to an SNS profile. This non-matched pair design mitigates the risk of detection, as our fictitious SNS profiles are very similar – even across gender conditions. In addition, to avoid revealing the experiment to subjects, we implement an extensive duplicate check to prevent the same advertiser from being treated more than once.[11]

In addition to reducing the risk of detection, a non-matched pair design reduces the costs to advertisers imposed by the experiment, since only one fictitious application needs to be reviewed. To limit unintended costs to advertisers, we reject all positive responses within 24 to 48 hours, noting that the fictitious applicant has already received a confirmation for another room at short notice.

To run the experiment, we code a computer program that performs randomization, collects data on room, apartment, advertiser, and roommate characteristics, and – after a final check by one of the authors – automatically applies for a vacant room. This reduces the risk of human error, collects data in a standardized way, and allows for a large number of observations by reducing costs. We then collect the responses and classify them as *callbacks* if the applicant is explicitly invited to an (online) visit or an (online) meeting with at least one of the potential roommates, as *other*, if the advertiser asks for additional information, or as *rejection* if the fictitious applicant receives a rejection for the room.

As additional control variables, we collect a variety of data on room and apartment characteristics, socio-demographic variables at the neighborhood and district level, i.e., data on number of residents, age of residents and fraction of foreign residents, Covid-19 data at the county level, advertiser data[12], text sentiment data, and profile statistics to determine how many people visit the profiles and view the photos.

2.4 Results

2.4.1 Descriptive Results

In professional markets, such as hiring, legislation in many countries prohibits discrimination based on gender or age. In contrast, vacant room ads often include age or gender restrictions. Regarding gender, about a quarter ($n = 866$) of the ads in our sample specify the gender of

[11] Sometimes, one advertiser places multiple ads for the same room or for a series of vacant rooms at the same time or in quick succession to increase visibility.

[12] Unfortunately, we do not have data on the nationality or ethnic origin of the advertisers. However, we do collect (a) the profile names of the advertisers and (b) information about the languages spoken by the incumbent inhabitants of the shared apartment. (a) Some profile names do not allow us to infer ethnicity (e.g., "sunflower91"). For others, we try to infer this information. When we control for the advertiser's alleged nationality in our regression models, we find no significant difference in callback rates with respect to the advertiser's alleged nationality. (b) The number or type of languages spoken also has no effect on callback rates. There were only a few cases ($n = 60$, 1.63 percent) where Turkish was among the languages spoken by the roommates – too few to run separate regressions.

the roommate sought. Since we generally adhere to these specifications, we do not randomly assign the gender of a fictitious applicant to ads that impose a specific gender restriction. In the following, we restrict our analysis to those ads where we randomly assign the gender of a fictitious applicant, i.e. we restrict our analysis to ads that do not impose gender restrictions ($n = 2,810$).[13]

Table 2.1: Mean Callback Rates

	% callback for German names	% callback for Turkish names	Ratio	%-difference (p-value)
All applications	42.54 [1,420]	36.62 [1,390]	1.16	-5.92*** (0.0013)
SNS-Profiles				
Without SNS-Profile	41.96 [684]	31.84 [691]	1.32	-10.12*** (0.0000)
With SNS-Profile	43.07 [736]	41.34 [699]	1.04	-1.73** (0.0436)
Females				
Without SNS-Profile	47.29 [332]	38.44 [346]	1.23	-8.85** (0.0200)
With SNS-Profile	48.61 [360]	45.77 [343]	1.06	-2.84 (0.4514)
Males				
Without SNS-Profile	36.93 [352]	25.22 [345]	1.46	-11.71*** (0.0008)
With SNS-Profile	37.77 [376]	37.08 [356]	1.02	-0.69 (0.8478)

Note: The table presents average callback rates for applicants with Turkish- and German-sounding names, respectively, using different subsamples. The numbers in brackets in each cell present the number of applications sent for the given subsample and ethnicity. Column 5 shows the p-values for a two-sample Wilcoxon rank-sum (Mann-Whitney) test of proportion testing the null hypothesis that the callback rates are equal across ethnic groups. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Of the 2,810 applications we send to ads that do not specify the gender of the potential roommate, we receive 1,434 (51.03%) responses.[14] The vast majority of these responses are callbacks ($n = 1,113$, 77.62%, see Table A.3, p. 153). 10.32% of the responses are rejections, and 12.06% are other responses, i.e., responses that ask for additional information. In our analysis, we focus on callbacks as the dependent variable, instead of "positive responses", which would include both, callbacks *and* other responses.[15]

Applicants with a German-sounding name who do not include a link to an SNS profile receive a callback for 41.96% of the applications they send, while applicants with a Turkish-sounding name without a link to an SNS profile that is apt to break prevailing ethnic stereotypes receive a callback only in 31.84% of the cases. Thus, minority applicants without an SNS profile must send 32% ($p = 0.0000$) more applications than majority applicants without an SNS profile to receive one callback (see column 4 of Table 2.1).

This difference between applicants with Turkish- as opposed to German-sounding names decreases to 4% ($p = 0.0436$) if the applications include a link to an SNS profile. Applicants

[13]As a robustness check, we rerun our main regression model (Table 2.2) for the full sample of ads (see Table A.4, p. 154) and also separately for those ads that specified the gender of the roommate sought (see Table A.5, p. 155). The coefficients of the Turkish dummy variable in the base specification (including control variables) are -0.101 ($p = 0.000$) for ads that do not specify gender, -0.120 ($p = 0.000$) for the full sample including all ads, and -0.144 ($p = 0.000$) for ads that accept only female or male applicants.

[14]For summary statistics on selected variables, see Table A.1 (p. 152).

[15]Running our main regression model using "positive responses" as the dependent variable as an additional robustness check, leads to similar results as our main regression model, see Table A.6 (p. 156).

with German- (Turkish-)sounding names that include a link to an SNS profile receive a callback in 43.07% (41.34%) of the cases. Thus, an SNS profile that is apt to break prevailing ethnic stereotypes about minority applicants helps ethnic minority applicants to close the gap with the ethnic majority applicants.

For female applicants, the difference between the ethnic groups is 23% ($p = 0.0200$) without an additional SNS profile included versus 6% ($p = 0.4514$) with an additional SNS profile included. The latter difference is no longer statistically significant. For male applicants, the difference decreases from 46% ($p = 0.0008$) to a negligible, statistically insignificant 2% ($p = 0.8478$), as reported in Table 2.1.

Turkish males are the most discriminated group. Without additional social media information that may break with prevailing ethnic stereotypes, Turkish-sounding male names receive callbacks only in 25.22% of the cases (see column 3 of Table 2.1). With additional social media information, the callback rate increases to 37.08%, which is very similar to male applicants with German-sounding names (37.77%) who include a link to a personal SNS profile. This negligible difference suggests that including social media information, which is apt to disrupt prevailing ethnic stereotypes, eliminates the ethnic gap for minority males.

Moreover, applicants with German-sounding names benefit hardly from the additional social media information. This is understandable, as the fictitious profiles may not be seen as informative when associated with a German-sounding name. For females with German-sounding names, the average callback rate increases by 1.32%, while for males it increases by only 0.84%. For applicants belonging to the majority group, we conclude that providing additional social media information does not contribute to a significant update of initially held beliefs.

For applicants with Turkish-sounding names, the increase in callbacks that resulting from the inclusion of an SNS profile is both statistically and economically significant: it amounts to +7.33% for females and +11.86% for males. That is, the additional SNS profiles appear to be highly informative for those advertisers that receive an application from a minority applicant, leading them to update their initial probabilistic beliefs about the likelihood that the applicant shares certain unobservable qualities. This is particularly true for the otherwise most disadvantaged (and likely most stereotyped) group: Turkish males.

2.4.2 Regression Analysis

In the following, we further investigate the interaction between ethnicity, gender, and social media information using different probit regression models to control for various variables that may affect the probability of receiving a callback. Our main specification is as follows. For a fictitious applicant i, the probability of a callback is given by:

$$\begin{aligned} Pr(Callback_i) = \beta_0 + \beta_1 * Turkish_i + \beta_2 * Female_i + \beta_3 * Instagram_i \\ + \beta_4 * Turkish_i * Instagram_i + \lambda^{'} \boldsymbol{X} + \epsilon_i, \end{aligned} \tag{2.1}$$

where Turkish is a dummy variable indicating whether applicant i has a Turkish- ($Turkish_i = 1$) or German-sounding name ($Turkish_i = 0$). Female is a dummy variable that equals one if applicant i is female ($Female_i = 1$) and zero if the applicant is male ($Female_i = 0$), and Instagram is a dummy variable that equals one if the application includes an SNS profile ($Instagram_i = 1$) and zero otherwise ($Instagram_i = 0$). This is followed by an interaction term of $Turkish_i$ and $Instagram_i$. $\beta_0, ..., \beta_4$ are unknown parameters, X is a vector representing different sets of control variables, ϵ_i is an error term. $Callback_i$ equals one if the respective applicant receives an invitation to visit the apartment ($Callback_i = 1$), and zero otherwise ($Callback_i = 0$), i.e., if applicant i receives a rejection, an other response asking for additional information, or no response at all.

Table 2.2 shows the average marginal effects of probit regression models using callback as dependent variable. The regressions include dummies for ethnicity, gender, and social media information, interactions between ethnicity and social media information, and a large set of control variables in columns 2-4. Standard errors are clustered at the city level.

The results indicate that having a Turkish-sounding name has a large negative and statistically significant effect on the probability of receiving a callback in all specifications. The estimated effect on the callback rate is -0.101 ($p = 0.000$) for Turkish applicants compared to German applicants (see column 2 of Table 2.2). Being female has a substantial positive and statistically significant effect on the probability of being called back.

The effect of including an SNS profile (Instagram) in the application is mainly driven by the applicant's ethnicity, as the interaction term of Turkish and Instagram positively affects the callback probability and is statistically significant at the 1% level (see column 2 of Table 2.2), which is consistent with the descriptive results presented above. Moreover, the effect sizes of Instagram are very similar to the ethnicity effect, confirming that social media information has the potential to resolve the ethnic bias.

In line with the descriptive results, the effect of providing additional social media information seems to be stronger for applicants with a male Turkish-sounding name (0.120 ($p = 0.009$) compared to 0.0604 ($p = 0.204$) for Turkish-sounding female names, see column 3 and 4 of Table 2.2). However, when testing for differences in the coefficients between the two subsamples, we cannot reject the equality of the coefficients of our main explanatory variables across gender. This underscores that additional social media information potentially eliminates ethnic discrimination – across *both* genders.

Table 2.2: Results – Probit Models (Average Marginal Effects)

Callback	(1) Full sample	(2) Full sample	(3) Females	(4) Males
Turkish	-0.105***	-0.101***	-0.0836**	-0.131***
	(0.0292)	(0.0216)	(0.0347)	(0.0237)
Female	0.107***	0.0963***	-	-
	(0.00978)	(0.0113)		
Instagram	0.0104	0.0205	0.0249	0.00907
	(0.0283)	(0.0261)	(0.0368)	(0.0328)
Turkish × Instagram	0.0881**	0.0837***	0.0604	0.120***
	(0.0392)	(0.0293)	(0.0476)	(0.0459)
Observations	2,810	2,550	1,254	1,296
City Control Variables	No	Yes	Yes	Yes
Room and Apartment Control Variables	No	Yes	Yes	Yes
Covid-19 Control Variables	No	Yes	Yes	Yes
District Demographic Control Variables	No	Yes	Yes	Yes
Time Control Variables	No	Yes	Yes	Yes
Pseudo R^2	0.0155	0.153	0.160	0.171

Note: The table reports the average marginal effects of probit regression models with callback as the dependent variable. The sample consists of vacant room ads that do not impose any restriction on an applicant's gender. Column 1 reports the main effects without control variables, Column 2 shows the main effects including controls. Columns 3 and 4 report the main effects including control variables, split by gender. Testing for differences between the coefficients of *Turkish*, *Instagram*, and *Turkish* × *Instagram* across the two subsamples in column (3) and (4), we cannot reject the hypothesis that the coefficients are equal. City variables include dummies for each city. Room and apartment variables include (among others) room size, rent, number and gender of incumbent roommates, distance to the city center, number of bus stops, train stations, and university buildings near the apartment, dummies for languages spoken by roommates, dummies for the type of cohabitation. Covid-19 variables include cases, deaths and ICU patients at the county-level. District demographic variables include population figures divided by total, foreign, and Turkish populations on district level. Time variables include dummies for each month in which the experiment took place. Standard errors (in parentheses) are clustered at the city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The results presented are robust to the inclusion of a large number of control variables on city, room, apartment, and neighborhood characteristics, as well as monthly dummies (demand for rooms in shared apartments is typically highest before the start of the semester, MLP Finanzberatung SE, 2021), and Covid-19 control variables. Furthermore, the effects are robust using different samples, specifications (see Tables A.4, p. 154, and A.9, p. 159) and dependent variables (see Tables A.6, p. 156, and A.8, p. 158).[16]

As an additional robustness check, we computed a probit model that also includes a three-way interaction of our main variables of interest, i.e., Turkish, female, and Instagram (see Table A.9, p. 159). Consistent with our main probit model (see Table 2.2), the results suggest that the effect of providing additional social media information is mainly driven by Turkish applicants, particularly Turkish male applicants, as the interaction effect of Turkish and Instagram is substantial and statistically significant at the 5% level, while the three-way interaction is not statistically significant (see column 4 of Table A.9, p. 159).

[16] On a more qualitative account, we also count the number of smileys in advertiser's responses and collect data on sentiment, i.e., the emotional state of the responses (see Table A.12, p. 162). We regress these data on several explanatory variables and find that, on average, applicants with female names and applicants who provide an SNS profile receive more smileys and more emotionally positive responses.

In line with the descriptive results, we further interpret the fact that the ethnic gap between applicants with Turkish- and German-sounding names, and Turkish male names in particular[17], is eliminated by adding social media information that is apt to break prevailing ethnic stereotypes, suggesting that statistical discrimination plays a significant role in explaining the ethnic bias.

2.4.3 SNS Profile Visits

To the best of our knowledge, our study is the first to collect SNS profile data on visits, impressions, and many other social media variables[18] and use these to test for a link between SNS profile visits and callback rates.

We manually collect aggregated profile statistics on a weekly basis.[19] While we do not know the exact proportion of roommates who actually visited the SNS profile (Instagram does not provide raw data on visits and impressions), we do know that, based on the average weekly visits per profile and the number of weekly applications sent using the respective name, each application generated, on average, about one visit to that profile (see column 4 of Table 2.3).

One social media application containing a Turkish- (German-) sounding name results in an average of 0.97 (1.04) visits to the respective SNS profile. With an average of 0.69 visits per application submitted, male profiles are generally visited much less frequently than female profiles (1.34 visits per application, see Table 2.3).[20] This trend can also be seen in Figure A.3 (p. 162), where male names in general and Turkish-sounding male names in particular are visited less frequently.[21]

Table 2.3: Estimated Average Profile Visits per Application

	German names	Turkish names	Total
Females	1.35 (0.82)	1.32 (0.91)	1.34 (0.87)
Males	0.75 (0.61)	0.64 (0.52)	0.69 (0.57)
Total	1.04 (0.78)	0.97 (0.81)	1.01 (0.79)

Note: The table reports means and standard deviations (in parentheses) of calculating the estimated average profile visits per application by computing the number of visits to a given SNS profile in a given week divided by the number of applications sent per name in the same week in the SNS condition.

Especially in the case of a male minority applicant, some advertisers may refuse to review additional social media information, i.e., to click on the provided link, after having read the applicant's name in the application text. This could be an indication of (some) taste-based

[17] In addition, having a Turkish-sounding male name significantly reduces the probability of receiving a rejection when the applicant includes a link to an SNS profile, see column 4 of Table A.7 (p. 157).

[18] See Table A.2 (p. 153) for summary statistics on social media variables.

[19] Since we receive 94.9% of all responses within a week of sending the application (70.3% within 24 hours), weekly visits should provide a valid estimate of the number of profile visits per application.

[20] The results of regressing the visit variables on various explanatory variables, such as the number of social media applications in a given week, ethnicity, and gender, point in a similar direction (see Table A.11, p. 161).

[21] As an additional robustness check, see Figures A.4 (p. 163) and A.5 (p. 163) for a replication of Figure A.3 (p. 162) using data on visits before and after the experiment to cover weekly visits without running the experiment.

discrimination against male applicants with Turkish-sounding names. Or the stereotypical perceptions of an advertiser or roommate may be so preconceived, e.g., due to persistent inaccurate statistical assumptions about group identity (Bohren et al., 2019), that additional SNS information no longer plays a role.

During the experiment, Instagram enabled the feature to access aggregated statistics at the post level, i.e., for individual or bundled images. Analyzing this data, we find that the weekly reach of non-subscribers decreases with the number of posts (see Figure A.6, p. 163), meaning that older posts are generally less likely to be viewed by advertisers. However, in contrast to Manant et al. (2019), who find that recruiters do not even perform one additional click on an SNS profile to view additional information, our data suggest that a large proportion of advertisers also appear to review postings that are comparatively old, i.e., a large proportion of advertisers screen almost the entire profile.[22] Manant et al. (2019) argue that search costs are the reason why recruiters do not screen the profile thoroughly. The difference with our results could be due to lower search costs on Instagram than, for example, Facebook (Manant et al., 2019), or that search costs play a smaller role in non-professional markets with close interaction of agents.

Estimating the effects of various social media variables on the likelihood of receiving a callback for different ethnic and gender subsamples (see Table A.10, p. 160) shows that most variables, such as number of subscriptions, likes, and impressions, have no or hardly any effect on the callback probability. Only the number of followers and the weekly reach have a small but statistically significant effect on the callback probability (see Table A.10, p. 160). Moreover, a higher number of followers/subscribers has a small positive and statistically significant effect on the callback probability in all specifications except the female subsample. This result suggests that a larger number of followers, "friends", or a larger social network in general can increase market success, i.e., callback rates. However, the effect sizes are rather small.

2.5 Conclusion

While several field experiments suggest that the provision of additional information can reduce discrimination (Bertrand & Duflo, 2017; Cui et al., 2020; Nunley et al., 2011), no study has yet analyzed the potential role of social media information in this context. This is despite the fact that information posted on SNSs has become increasingly important over the past decade (Acquisti et al., 2015), with a significant proportion of employers using SNSs to scout for potential hires (Becton et al., 2019; Rosen et al., 2018) or to determine the applicants' hirability (Roth et al., 2016).

In our study, we investigate whether and to what extent social media information is able to reduce (statistical) discrimination by inducing an update of initially held conditional probabilistic beliefs. We apply to 3,676 vacant room ads in the fifteen largest student cities in Germany using fictitious identities. We randomly assign either Turkish- or German-sounding, female or male names to the applications. Except for the names of the profile owners, the four male profiles and the four female profiles are identical. The only difference between the profiles of one gender

[22]Since data access was enabled during the experiment, past data is only available retrospectively for postings 1-22, see Figure A.6 (p. 163).

is the name assigned to them: two profiles of each gender are assigned a Turkish-sounding name, the other two are assigned a German-sounding name. Each application is randomly assigned one of the eight names and whether it contains a link to the corresponding Instagram profile.

The results of our field experiment indicate that social media information, when appropriately designed, significantly reduces ethnic discrimination in selection. More specifically, the social media information in our study portrays what would appear to be an "average" student living in Germany, and are thus apt to break prevailing ethnic stereotypes when combined with an ethnic minority name. Thus, the social media information may lead the uninformed side of the market to update its potentially stereotyped initial conditional probabilistic beliefs about ethnic minority applicants.

We find that social media information is assessed and exploited by market participants in such a way that its provision strongly affects selection decisions. In contrast to brief self-descriptions by applicants themselves (Moritz & Manger, 2022), providing additional social media information hardly increases the probability of receiving a callback for applicants with German-sounding names, while it significantly increases for applicants with Turkish-sounding names. Moreover, the ethnic gap between minority and majority applicants disappears when their applications include a link to an SNS profile that is apt to break ethnic stereotypes. The most discriminated group, male applicants with Turkish-sounding names, benefit the most from the additional social media information provided. This makes our study one of the few to investigate mechanisms that have the potential to reduce or even eliminate (statistical) discrimination (Bertrand & Duflo, 2017; Zschirnt & Ruedin, 2016).

Moreover, our study is the first to investigate the impact of social media information on callback probability in a non-professional market. We are the first to include easily accessible social media information in the application text, thus reducing search costs to a minimum. At the same time, an applicant's ethnicity and gender are already indicated in the application text. Consequently, unlike previous studies, we can isolate the effect of social media information and analyze the extent to which it is contingent on an applicant's ethnicity and gender – unlike previous studies. We thus extend the sparse experimental literature on ethnic discrimination and the role of SNS profiles that provide information on otherwise unobserved applicant qualities. In addition, our study is the first to analyze extensive data on profile statistics, including profile visits and image views.

In terms of profile statistics, we find that an application that includes an SNS profile generates, on average, about one visit to the respective profile. Furthermore, image-level data indicate that the majority of advertisers screen almost the entire profile, suggesting that advertisers in non-professional markets review SNS profiles thoroughly. This finding contradicts the results of Manant et al. (2019) in hiring.

We interpret our results as primarily belief-driven statistical discrimination because social media information has the ability to challenge initial beliefs about unknown characteristics of a minority group applicant. Thus, by providing an SNS profile, imperfect information about otherwise unobservable applicant qualities is resolved and, if the information provided is apt to break prevailing ethnic stereotypes, the ethnic gap between minority and majority applicants may be dissolved.

One limitation to the generalizability of our findings refers to the way in which the fictitious SNS profiles were created. For our purpose, we deliberately designed the SNS profiles in such a way that, if included in the application of an ethnic minority applicant, they might break with prevailing ethnic stereotypes, leading to a (favorable) update of advertisers' initially held probabilistic beliefs. Including an SNS profile that instead reinforces rather than breaks prevailing ethnic stereotypes, will most likely not positively or even negatively affect the callback probability of ethnic minority applicants. Future studies may want to vary the content of SNS profiles to improve our understanding of *how* the information provided on SNSs affects the conditional probabilistic beliefs of the uninformed side of the market.

Furthermore, our study is embedded in a specific setting, the shared housing market in Germany, and it refers to two specific ethnic groups in this setting (applicants with German- versus Turkish-sounding names). While we are confident that our results are generalizable to other markets plagued by asymmetric information, future studies might analyze whether and to what extent social media information can reduce (statistical) discrimination in other markets and contexts.

Chapter 3

Visual Stereotypes & Inequality in Informal Networks and Markets[1]

3.1 Introduction

Beliefs and attitudes toward members of minority groups are a major barrier to their access to social networks and employment, housing and other markets – leading to large and persistent inequalities. Economic theories suggest that economic decisions, which affect market access, are essentially determined by choices that mirror preferences, constraints, and beliefs. However, beliefs can be influenced by (visual) narratives in mass and social media (Ash et al., 2021; Gehring et al., 2022; Wittenbrink & Henly, 1996) potentially leading to stereotypical and inaccurate beliefs (Bohren et al., 2023; Bohren et al., 2019; Bordalo et al., 2016) about individuals with certain characteristics, such as ethnicity, gender, or others (Coffman et al., 2023).

Social media has dramatically changed the way we communicate, interact and consume information – fundamentally transforming our daily lives (Allcott et al., 2020). Moreover, social media offers a way to access otherwise inaccessible information about, e.g., a wanna-be friend, job applicant, potential roommate, or tenant (Acquisti & Fong, 2020; Galos, 2023; Kluemper et al., 2012). Therefore, it is important to understand how information from social media shape beliefs and influence decision making about participation of disadvantaged groups.

In this paper, we analyze the role of visual narratives on selection decisions and participation using a combination of experimental evidence on how ethnic minority status and visual stereotypes affect, (1) participation in social networks and (2) the housing market. To the best of our knowledge, this paper is the first to investigate the role of visual stereotypes on discrimination and market outcomes for disadvantaged groups. Moreover, we study our research question in two different informal settings: Personal social network formation and the shared housing market. In

[1]Chapter 3 is based on the working paper "The Visual Narrative: Stereotypes, Unequal Treatment, and Informal Networks" by Raphael Moritz (University of Tübingen), Christian Manger (University of Tübingen), and Kerstin Pull (University of Tübingen). The implementation of the field experiments was approved by the Ethics Committee of the Faculty of Economics and Social Sciences of the University of Tübingen (A2.5.4-286_bi) on May 2, 2023 and April 24, 2024 (A2.5.4-286.1_hb). The research project was pre-registered in the AEA RCT Registry (#11322 and #13574): https://doi.org/10.1257/rct.11322-3.0 and https://doi.org/10.1257/rct.13574-1.0.

doing so, we bridge the gap between online self-representation and offline economic outcomes, demonstrating how digital identities can affect real life inequality.

We conduct two pre-registered field experiments using fictitious Instagram profiles that varied systematically by ethnicity (minority vs. majority name) and the presence of visual narratives, i.e., stereotypical images associated with cultural identification and religious affiliation of the ethnic minority. Half of the profiles conforms to existing ethnic minority stereotypes, whereas the other half does not. We validate the design of the experimental conditions and visual stereotypes through multiple online experiments. In these, we find that participants are able to successfully attribute the intended ethnicity to the fictitious applicant, and that the photos unambiguously and precisely signal commonly held ethnic minority stereotypes. Also, the social media profiles are not perceived to be fake.

Despite the fact that social media represents one of the primary channels for misinformation (Allcott & Gentzkow, 2017), often targeting ethnic minorities and/or immigrants (Ekman, 2019), there is a paucity of empirical evidence on the role of specific types of social media contents on the formation of ethnic stereotypes and the selection outcomes of ethnic minority applicants.[2]

This paper adds to the literature on the role of social media information affecting selection decisions (see Chapter 2 of this thesis; Acquisti & Fong, 2020; Manant et al., 2019; Nüss, 2024). However, none of these studies have examined the role of stereotypes and how different social media *contents* affect selection decisions. Moreover, we investigate unequal treatment in both social (friends) and more economic (roommates) domains, facilitating a comprehensive understanding of whether unequal treatment vary across informal online and offline contexts.

Moreover, understanding the formation of informal social networks is critical as larger social networks are associated with higher levels of life satisfaction and perceived social support (Manago et al., 2012). Furthermore, larger networks and audiences facilitate the dissemination of information (Bond et al., 2012), job search (Dustmann et al., 2015), or promotion (Zinovyeva & Bagues, 2015) – with persistent and substantial effects contributing to long-lasting inequality (Ioannides & Datcher Loury, 2004).

Existing experimental literature concentrates only on professional social networks, such as job networks and their information provision (Evsyukova et al., 2025) or the formation of academic networks (Ajzenman et al., 2025). However, none of the studies investigated *non-professional* social network formation – which may affect not only job search, but also other aspects of social life, including searching for a room, an apartment, or even friendship, love, attention, and social affiliation in general. In addition, no previous study has examined the effects of varying stereotypical information on social network formation. Our paper fills these research gaps by investigating how ethnic minority status, cultural identification, and religious affiliation affect the formation of informal social networks, and, more generally, how these factors contribute to inequality in non-professional, informal domains.

In the first experiment, we send more than 1,000 friend requests to potential friends suggested by the social media platform Instagram. Each potential friend of anyone of our four

[2] For a more general overview of the experimental literature on unequal treatment, see Riach and Rich (2004), Bertrand and Duflo (2017), and Lippens et al. (2023).

fictitious accounts receives a request from either a minority or majority profile, with or without stereotypical photos. We measure whether these requests are accepted or not and whether subsequently, potential friends subscribe to our profile themselves in order to causally determine whether ethnicity and/or visual stereotypes affect informal social network formation. We collect all publicly available information about the potential friends to be able to analyze potential determinants of discrimination.

Overall, we find a 46% lower friend acceptance rate for minority profiles (28%) compared to majority profiles (41%). Examining whether connections subsequently add our fictitious profile as a friend on their own profile after accepting our friend request, we do not find a significant difference. The “re-follow” rate of minority profiles (11%) is similar to the one of majority profiles (12%).

However, showing stereotypical photos on one’s profile has a significant negative effect on both, friend acceptance rates and re-follow rates. The acceptance rate of minority profiles with stereotypical photos is 69% lower than the one of majority profiles, and the re-follow rate is 78% lower. Interestingly, minority profiles without stereotypical photos receive more re-follow requests than majority profiles.

In the second experiment, we use the very same social media profiles to apply for approximately 3,100 vacant room ads on the largest platform for shared apartment ads in Germany to causally investigate the effect of visual stereotypes on economic decision making and market access beyond social networks. In Germany, about one third of students live in shared apartments (Kroher et al., 2023). Furthermore, it is common to include a social media profile when applying for a room, and evaluating social media information does not violate anti-discrimination laws (see Chapter 2 of this thesis).

The results of the second experiment reveal an overall 61% lower callback rate for minority applicants (20%) compared to majority applicants (32%). Linking a social media profile without stereotypical photos slightly reduces the gap to 44%, while minority applicants with a profile that includes stereotypical photos are 71% less likely to be called back. Across all treatments, the inclusion of stereotypical content negatively affects callback rates for both ethnic groups, but the effect is particularly detrimental for minority applicants.

Both experiments reveal significant discrimination against individuals with ethnic minority names, especially when minority applicants reveal cultural identification and religious affiliation via stereotypical photos. The ethnic gap is larger in access to offline vacant rooms compared to online social networks. Thus, access to social networks is easier for ethnic minority members compared to access to housing. Since the costs of accepting a virtual friend (and possibly later dissolving it) are much lower than inviting a potential roommate in one’s home, it is plausible that we observe less discrimination.

In general, our results indicate that visual narratives significantly affect the selection behavior of market participants, and thus may hinder participation in online social networks, as well as offline housing or other rather informal markets. The results show that the persistent disadvantage faced by minority applicants and the exacerbating effect of stereotypes call for increased awareness and interventions to address implicit bias in online and offline contexts and especially in informal, non-professional economic and social settings.

The remainder of this chapter is organized as follows: Section 3.2 reviews different literature that our analysis speak to and highlights our contribution to those literature. Section 3.3 describes the experimental setup of our two experiments. Subsequently, we present the results of our experiments in Section 3.4. Section 3.5 presents a short summary of the study's limitations, and Section 3.6 concludes.

3.2 Literature and Contribution

The following chapter describes the literature on how stereotypes affect both social and economic behaviors, with a particular emphasis on the visual representation of stereotypes (Section 3.2.1), the role of social media platforms (Section 3.2.2), and discriminatory behavior and the role of social media platforms (Section 3.2.3). Building on foundational work in economics and social psychology, we highlight how statistical discrimination and cognitive shortcuts can perpetuate inequality, particularly when (visual) stereotypes reinforce existing biases. However, despite substantial research on social media and on discrimination, the role of visual stereotypes in affecting real-world outcomes for ethnic minority groups remains underexplored.

3.2.1 Stereotypes

Economists have contributed significantly to understanding how stereotypes affect markets and behaviors. Arrow (1973) and Phelps (1972) introduced the concept of statistical discrimination, where employers or other decision-makers use group averages as proxies for individual attributes in the case of asymmetric information. While this may be economically rational, it perpetuates inequality by penalizing individuals based on group membership rather than personal merit. Recent literature explicitly takes account of the fact that statistical discrimination by market participants may result from actually inaccurate beliefs (Bohren et al., 2023) and that stereotypes may in fact be based on erroneous assumptions (Bordalo et al., 2016).

In social psychology, stereotypes are defined as "mental representations of real differences between groups" that allow "easier and more efficient processing of information about others" (Hilton & von Hippel, 1996, p. 240-241). Thus, stereotypes serve as cognitive shortcuts that highlight distinctive traits of groups while neglecting common characteristics, enabling individuals to process vast amounts of information efficiently (Bordalo et al., 2016; Hilton & von Hippel, 1996). Stereotyped beliefs may lead to inaccurate judgments, resulting in extreme and unlikely representations. Turner et al. (1979) introduced social identity theory, positing that people categorize others into in-groups and out-groups. This categorization lays the foundation for stereotypes, as individuals ascribe group-specific characteristics based on limited information. The representativeness heuristic described by Tversky and Kahneman (1983) also provides a cognitive underpinning for stereotypes. Their concept demonstrates how people rely on prototypical examples to make judgments, even when these examples do not reflect actual probabilities.

As the formation, maintenance, application, and change of stereotypes are influenced by both cognitive and motivational factors (Hilton & von Hippel, 1996), stereotypes or stereotyped beliefs may be subject to changes if new information alters a group's distinctive traits (Bordalo et al., 2016; Coffman et al., 2023). Yet, challenging stereotypes needs cognitive control – which depends on motivation and cognitive resources as stereotypes are often triggered automatically (Devine, 1989).

However, while stereotypes are often deeply ingrained and persistent over time, interventions can reduce their impact. Bohren et al. (2019) show that providing positive evaluations about disadvantaged groups as a result of biased beliefs can mitigate unequal treatment against women. Their study finds that individuals correct their biases when presented a sequence of positive evaluations. In social psychology, J. C. Becker and Swim (2011) highlight the effectiveness of counter-stereotypical exemplars in changing perceptions.

The literature on visual representation of stereotypes is scarce and focuses particularly on stereotypes in online gaming and news reporting. These visual biases affect perception and decision-making, with individuals being more likely to identify objects or traits that align with racial or gender stereotypes (Correll et al., 2015). Moreover, visual elements in media have been shown to leave a stronger impression on memory compared to auditory information, reinforcing their impact (Graber, 1990). More recent research emphasizes the pervasiveness of visual stereotypes. Ash et al. (2021) find that men and white individuals are frequently overrepresented in news media, while women and ethnic minorities are often underrepresented or they are depicted in stereotypical roles, reinforcing the need for critical evaluation of media content.

However, the role of visual stereotypes and its implications for ethnic minorities and their economic opportunities is yet unknown. We fill this research gap by causally identifying how visual representations of different kinds of stereotypes on social media affect the formation of (online) informal social networks and (offline) economic outcomes in the search for a shared apartment. By bridging the domains of online self-representation and offline consequences, we address a critical gap in understanding how visual information that represent commonly held stereotypes shape both social and market outcomes.

3.2.2 Social Media Platforms

The advent of social media and digital platforms in general has introduced new dimensions of studying ethnic discrimination and the role of stereotypes in decision-making. Online platforms both reflect and reinforce stereotypes by amplifying biases in algorithms and user-generated contents (Bailey et al., 2013; Singh et al., 2020). Additionally, social media platforms are an important source of information about individuals and provide reliable personality assessments that correlate with job performance, hirability, and academic performance, providing valuable information for employers, landlords, colleagues, friends, or others (Kluemper et al., 2012).

A large part of the economic literature focuses on the effect of information on social media platforms or exposure to social media on misinformation and the results of political elections (Allcott & Gentzkow, 2017; Aridor et al., 2024; Guess et al., 2023). The economics of social media encompasses content production, distribution, and consumption, with implications for misinformation, segregation, and political outcomes (Aridor et al., 2024).

Allcott et al. (2020) finds that deactivating Facebook can reduce political polarization and increase subjective well-being. Exposure to counter-attitudinal news on social media can decrease negative attitudes towards opposing political parties (Levy, 2021). Furthermore, social media algorithms may limit exposure to counter-attitudinal content, potentially exacerbating polarization and stereotypes (Levy, 2021). Allcott and Gentzkow (2017) demonstrate that misinformation on social media often exploits stereotypes to increase engagement, with the potential to entrench societal divisions.

The previous literature shows that social media has an important effect on information consumption and the formation, prevalence and evolution of stereotypes. As the contents presented on social media platforms changed from text-based information to more graphical contents, such as photos or videos, in the last decade (see Chapter 2; Aridor et al., 2024), understanding the role of visual representations of stereotypes and how they affect market outcomes of ethnic minority members is an important research gap that we address in our study.

3.2.3 Discrimination and Social Media

Social media platforms can both perpetuate and provide grounds for discriminatory practices. In recent years, there has been a notable increase in experimental evidence examining the role of social media platforms and ethnic discrimination.[3] For example, Gunarathne et al. (2022) presents evidence on ethnic discrimination with customer complaints, finding that African-American customers are less likely to receive responses to complaints to major U.S. airlines on Twitter.

Two studies explore ethnic discrimination in the formation of *professional* social networks. Ajzenman et al. (2025) create fictitious Twitter accounts of PhD students in economics, finding that ethnic majority accounts received 12% more follow-backs than ethnic minority accounts. Notably, African-Americans from top universities do not receive more follow-backs compared to Whites from lower-ranked institutions. Similarly, Evsyukova et al. (2025) used Linkedin profiles to investigate racial discrimination in job networks. They find that connection requests from African-American profiles were 13% less likely to be accepted. The racial gap resolves when contacting the connections for career advice.

Both studies highlight significant disparities in the formation of professional social networks for ethnic minority groups. However, no prior research has yet examined how (variations in) stereotypical information affect social network formation, nor has discrimination in non-professional social networks been investigated.

Manant et al. (2019) and Acquisti and Fong (2020) focus on social media information and ethnic discrimination in the labor market, while in Chapter 2, we conduct a similar experiment in the housing market which has the important advantage that personal social media information can be integrated directly into the application, eliminating search costs and ensuring that recipients can exploit the information (see Section 2.1 of this thesis).

Manant et al. (2019) conduct a field experiment in the French job market and find that recruiters use information from Facebook profiles to discriminate against ethnic minority applicants, with a 42% gap in callback rates. However, this ethnic gap disappeared when an exogenous

[3] Apart from *ethnic* discrimination, Baert (2018) finds that job applicants with less beneficial Facebook profile pictures are significantly less likely to be called back. Nüss (2024) investigates hiring discrimination against union members, revealing that job applicants expressing pro-union sentiment on social media receive less callbacks.

profile layout change (on Facebook) reduced the salience of ethnic information. In a similar experiment, Acquisti and Fong (2020) investigate the effect of religious and sexual orientation information from social media profiles on the callback rate of applicants in the U.S. While they find no significant national-level bias, employers in areas with higher shares of Republican voters show significant bias against Muslim candidates compared to Christian candidates. The study, that is closest to our research design, is the one by Moritz et al. (2023) who demonstrate that providing social media profiles that challenge ethnic stereotypes as additional information within an application for a vacant room can almost eliminate discrimination against ethnic minority applicants (see Chapter 2 of this thesis).

However, none of the studies above investigate the role of (a variation of) stereotypical information on ethnic discrimination. Furthermore, we are the first to present evidence on the interplay between visual stereotypes, digital identities and offline opportunities by exploring the discourse on discrimination and inequality in interconnected social and economic settings.

3.3 Experimental Design

We study the effect of ethnicity and visual stereotypes on social media profiles on informal social network formation (Study I) and callback rates in the shared housing market (Study II). We begin by setting up social media profiles on Instagram and creating stereotypical images using data on common ethnic minority stereotypes held by the ethnic majority. To signal ethnicity, we use distinctive Turkish and German-sounding names for the ethnic minority and majority, respectively. In Germany, about one third of the population has a migration background, with people of Turkish origin being the largest group of ethnic minorities in the country (Federal Statistical Office, 2022a, 2022b). In what follows, we describe our experimental procedure.

3.3.1 Creating Realistic Instagram Profiles

To investigate the effect of visual stereotypes on ethnic discrimination, we create a total of four fictitious Instagram profiles. Instagram is the most used social media platform in Germany (Koch, 2023). The platform focuses on photo and video sharing. Registered users can upload images and videos, add built-in filters, descriptions, hashtags, geographical locations and share them with other users. Profiles and their contents can either be shared publicly with everyone or only with (previously accepted) friends. Users can explore content shared by others, engage by liking or commenting on photos, and follow other users.

Instagram was first launched in 2010 and acquired by Facebook (now “Meta Platforms”) in 2012. According to advertising statistics, the social network has approximately 1.6 billion users around the world, indicating that 32.2% of the world’s population aged 13 and above use Instagram (Datareportal, 2024)[4]. Instagram has 33.8 million users in Germany, the majority of whom are under 30 years old (Datareportal, 2024; Koch, 2023).

[4]This figure excludes Chinese internet users, as Instagram is blocked in China (Datareportal, 2024).

Our profiles show a 24-year-old male student[5] who – according to a pretest for Chapter 2 (see Section A.5.1, p. 163) – might be of German or Turkish origin. The profiles were created in July and August 2019 and feature images of the profile owner alone ("selfies") or with friends, images of food, nature, and photos from activities, such as traveling – including a story highlight from the alleged hometown of the profile owner.[6] All images are consistent with the types of images frequently posted on Instagram and with data collected from existing Instagram profiles of business administration students from Germany (see Chapter 2; Hu et al., 2014). In line with Chapter 2, we add short captions including hashtags written by student assistants who use Instagram frequently.

The names of our fictitious applicants are based on Chapter 2 which uses most popular Turkish and German names of the respective birth cohorts from which we choose "Muhammed Kaya" as an ethnic minority name and "Tobias Weber" as an ethnic majority name, since these names show the largest effects of additional information from personal social media profiles (see Chapter 2 of this thesis). In a randomized online pilot experiment, we examine whether participants can accurately identify each name's intended ethnicity (see Section B.6.1, p. 219). This first online pilot experiment was conducted in June 2023 using a professional survey platform, recruiting $n = 1,725$ students. The results reveal that 91.9% of respondents indicate that Muhammed Kaya is likely or very likely to be a person of Turkish origin, while 97.8% of respondents indicate that Tobias Weber is likely or very likely to be a person of German origin.

3.3.1.1 Visualizing Stereotypes

In the first pilot experiment (see above), we also ask participants which characteristics they typically attribute to ethnic minority members to create social media profiles that either *contradict* or *conform* to ethnic minority stereotypes (see Section B.6.1, p. 219). Participants were randomly assigned to ethnic minority or majority conditions and asked to list characteristics and attributes that they associate with people of Turkish or German origin in open-ended questions. We evaluate the open-ended responses using content analysis, which reveals patterns of perception such as Turkish names being associated with patriotism, religiosity, and tradition, while German names were linked to punctuality, discipline, and environmental consciousness.

In line with the results and previous literature on Turkish stereotypes held by Germans (Baur & Ossenberg, 2016; Ossenberg, 2019), we searched and collected the contents publicly available on Instagram that are posted by male profile owners with Turkish migration background regarding religiosity, patriotism, attachment to tradition, and national pride. In addition to these, we select common themes posted on social media by young men with Turkish migration background living in Germany, such as smoking Shisha with male friends or the profile owner with a strongly motorized car. We record captions, locations, and hashtags which were indicated on such images that were posted frequently (see Figure B.10, p. 217) and which mirror

[5] 24 years is the average age of students in Germany (Federal Statistical Office, 2022d). In contrast to Chapter 2, we focus exclusively on males to limit the number of treatment arms.

[6] "Stories" are a function on Instagram that display full-screen images that can only be viewed for a short period of time. The "Highlights"-function allows users to display them permanently on their profile. See: https://about.instagram.com/de-de/blog/announcements/introducing-stories-highlights-and-stories-archive [Retrieved: April 16, 2024].

common Turkish stereotypes held by Germans (Baur & Ossenberg, 2016; Ossenberg, 2019) to construct representative images in line with previous research (De Leyn, 2023).[7]

To half of the Instagram accounts we add Turkish stereotypical images, showing the profile owner while visiting a Turkish mosque, smoking shisha with (male) friends in a shisha cafe, with a sports car, and with a flag of Turkey.[8] The images are accompanied by brief captions in German or Turkish incorporating hashtags and symbols/emojis which align with the content and captions of similar posts.[9]

In a second randomized online pilot experiment (between-subjects design), we ask students ($n = 506$) to rate and assess the resulting fictitious social media profiles in order to check whether the treatment conditions successfully manipulate relevant beliefs (see Section B.6.2, p. 225). The results show that profile owners of social media profiles that contain Turkish stereotypical images are significantly more often perceived as being patriotic, religious, and traditional, compared to the profiles without these additional stereotypical images. Moreover, the respective profile owners were significantly more likely to be regarded as being of Turkish origin.[10]

In addition, we asked participants of our second pilot experiment to assess the authenticity of the fictitious Instagram profiles and ask for comments on their validity. The profiles were perceived to be highly realistic and authentic in the vast majority of cases (see Tables B.35, p. 231, and B.39, p. 236): 85.1% of participants indicated that the social media profiles without ethnic minority stereotypes were extremely or very realistic, while for the profiles with images matching minority stereotypes, this was the case for 87.4% of participants.

3.3.1.2 Social Network Creation

In contrast to Evsyukova et al. (2025), who start their experiment with no prior connections, we opt to create an initial network of friends that is demographically similar to the profile owner. This approach offers several advantages: Firstly, having a pre-existing social network, even a relatively small one, reduces the likelihood of the profiles being perceived as inauthentic or fake.[11] Secondly, it mitigates potential endogeneity concerns. The likelihood of a user accepting a connection request may increase with the number of connections the requesting profile already has, as users might evaluate the legitimacy of connection requests based on visible social network size – presumably with a lower bound. Profiles with few or no connections may raise suspicions of being inactive or fabricated.

To establish an initial social network of friends for the fictitious profiles, we recruit students to subscribe to the profiles and engage with their content by liking a random selection of photos.[12]

[7] For a discussion on the ethical implications of our research, see Section B.8 (p. 236).

[8] In addition, we upload a so-called "Reel", which is a short video showing the city of Istanbul, Turkey. See Figures B.8a (p. 213), B.8b (p. 214), B.9a (p. 215), and B.9b (p. 216) for screenshots of the social media profiles.

[9] For instance, the photo in the mosque was accompanied by the following caption: "In the name of Allah, the most merciful #verse #happyfriday #friday #discover #prayer" (translated from Turkish).

[10] A majority applicant who displays Turkish stereotypes on his social media profile might actually be a second-generation immigrant – a German-born child of one foreign-born parent – and thus might already possess a German first and last name.

[11] Ajzenman et al. (2025) asked approximately 30 economists and students to follow their fictitious Twitter profiles to enhance credibility.

[12] To enhance realism, we adjusted the randomization process to prioritize newer photos, thus mimicking a profile's evolution over time.

For this purpose, we conduct an additional randomized online survey ($n = 528$) in order to create a realistic and representative follower base (see Section B.7, p. 234). We conducted the online survey in three waves between January and July 2023 using a professional survey platform. Each participant was randomly assigned to a profile and asked to subscribe and engage as described above. By randomly assigning participants to profiles, the characteristics of subscribers, such as their demographics or behavior, are distributed more evenly across profiles – which is an important prerequisite for Study I (see Section 3.3.2, p. 40).

The process of creating an initial social network ensures a realistic and representative follower base for the experimental profiles, enhancing their credibility and aligning with the demographic characteristics of the fictitious profile owner. Additionally, we followed back the "new" friends in order to create realistic friends networks. Participants were asked to remain subscribed for the duration of the experiments.

Furthermore, some participants were asked to subscribe to additional private accounts. In addition to the four fictitious social media accounts, we created a set of additional fictitious profiles representing students who are similar to our fictitious applicants. We use these additional, non-public profiles as fictitious friends and link to them on some of the photos of our fictitious accounts where friends can be seen.

3.3.1.3 Additional Profile Features

The mean number of likes on all posts on the profiles is about 1,176, while older posts have fewer likes than newer posts, as described above. The profiles have an average of approximately 228 subscribers and 263 subscriptions. The majority of subscriptions are followers that have subscribed to our profiles (see Section B.7, p. 234), and some public profiles that are frequently followed by students, such as travel, news, or fun profiles. We do not subscribe to profiles that post stereotypical images or any kind of religious or patriotic contents.

The majority of friends/subscribers, if indicated, live in Tübingen (36.5 – 40.7%), followed by the cities Stuttgart, Munich, and Reutlingen. The majority (62.1 – 72.2%) identify as female and are below the age of 34 (91.1 – 93.6%).

The two profiles without stereotypical images feature 27 posts, while the two profiles with stereotypical images feature 32 posts. All profiles we use for the experiment are set up as public profiles so that most of the information (name, profile picture, biographical information, and posts) is available to everyone, allowing anyone to view the profiles at basically no cost.[13] We add a short bio (biographical information) which is equal across profiles and contains three pieces of information: age (24), occupation (student, together with an emoji symbolizing a graduation cap), and location (Tübingen).[14]

[13]Non-registered users can view the full profile. However, a login is required to access additional information, such as additional images within a post or the list of friends/subscribers.

[14]The "bio" section can be found at the top of a user's profile page, where users may add a personal description of up to 150 characters.

3.3.2 Study I: Formation of Informal Social Networks

Our first pre-registered field experiment (AEA RCT Registry ID #13574) investigates the role of ethnicity and visual stereotypes on the formation of informal social networks. Conducted on Instagram between May and June 2024, the field experiment involved sending friend requests to $n = 1,061$ individuals recommended by the platform's algorithm. The experiment employs a balanced 2x2 design, with half of the friend requests originating from profiles with ethnic minority (majority) names and half of the profiles exhibiting visual stereotypes, while the other half does not.

In a first step, we collect potential subjects (friend suggestions) for every of the four fictitious social media profiles from the platform (see Figure B.11, p. 217, for a screenshot on the suggestion function of Instagram). In cases where a subject is suggested more than once, we randomly select one of the four treatment conditions, i.e., profiles, to send a friend request to the respective user. Each suggested friend receives one request to minimize the risk of detection.

Suggestions are primarily based on the existing social network of the respective profile. As the existing network of the profiles primarily consists of students who are randomly selected to be part of the network of a particular profile, the sample selection is arguably quasi-random due to the random assignment of participants in the additional online survey that we used to create the initial network (see Section 3.3.1.2, p. 38).[15] Furthermore, if participants who self-select into the survey were more conscious of discrimination, we would observe a rather lower bound of discrimination in our field experiment. This would be even more prevalent if the social networks of these individuals also exhibited high awareness of unequal treatment.

To prevent subjects familiar with the experiment or associated profiles from being treated, we compile a comprehensive list of usernames, including all subscribers and subscriptions across all profiles associated with the experiment (including those that are not part of the study) and all participants from all pilot studies. These usernames are added to a block list to ensure they are excluded from the subject pools. This minimizes the risk of contamination by preventing individuals with prior knowledge of the experiment from affecting the results or inadvertently revealing the purpose of the study.

We excluded verified profiles from the subject pools, as these profiles mainly belong to public figures, celebrities, and/or influencers who typically do not react to friend requests. We did not exclude business or non-profit profiles as only a very small number of business or non-profit profiles were suggested as potential connections.

Both private and public profiles are included as potential subjects. The type of profile has important implications for the experimental outcomes: Private profiles require explicit approval (or decline) for each friend request, while in the case of public profiles, new friends/subscribers are automatically added to their list.[16] Thus, users of private profiles can more actively manage

[15] A small feasibility study revealed that the pool of suggested friends on Instagram remains largely stable when users do not actually send friend requests to the suggested accounts. Thus, in the case of a random assignment of suggestions from all treatment accounts, the pool of potential subjects of each account would have been significantly reduced for each subsequent round, rendering the experiment unfeasible. We address potential endogeneity in the later analysis (see Section 3.4.1.4, p. 48).

[16] However, the owners of public profiles have the option to manually remove subscribers if desired.

their network. Both private and public profiles, however, retain the capability to re-follow our profiles after our request.

To conduct the experiment, we write a computer program that manages the subject pools, by collecting friend suggestions and checking their eligibility (as described above) and the sending of friend requests. Then, the program collects all publicly available data from the profile, such as name, biographical information, number of posts, subscribers, subscriptions, mutual connections (if any), joining date, and number of username changes, and stores them in a database. With this information, we can explore heterogeneity in unequal treatment based on user characteristics. Importantly, we do not collect any ex-post information that only becomes available after a particular (private) user accepted the friend request.

When subjects receive a friend request, they are notified showing the profile picture, the username, and the name of the requesting profile.[17] As both, the username and the name signal the ethnicity of the profile owner, the ethnicity signal is very salient. As described above, we did not vary the profile pictures. Thus, the only variations are the ethnicity signal and the visual stereotypes. All other information is held constant.

Throughout the experiment, we regularly check whether our profiles are "flagged for review" – a feature of Instagram to detect spam or fake accounts. However, this was never the case.

After sending friend requests, the program frequently revisits the profiles that received a friend request to monitor their response behavior, if any. For private profiles, it determines whether the user has *accepted* or *rejected* the friend request. For both private and public profiles, it checks if the user has (also) *re-followed* the experimental profile. If a user reacts in any way, the program stores the type of reaction and again fetches all mutual connections. This step ensures the ability to analyze a potential source of endogeneity, as a user may decide to accept a request, for instance, because another mutual connection has already done so in the meantime (see Section 3.4.1.4, p. 48).

In addition, we manually collect data on profile visits, impressions, and other engagement metrics for every profile throughout the experiment.

3.3.3 Study II: Inequality in the Housing Market

In our second pre-registered field experiment (AEA RCT Registry ID #11322), we examine how visual stereotypes influence selection decisions, focusing on the impact of online self-presentation on offline economic outcomes. By employing a between-subjects, non-matched correspondence test design, we illustrate how digital identities can shape economic opportunities. The study is embedded in the market for shared apartments and is conducted on Germany's largest platform for shared apartment advertisements, focusing on long-term rentals.

The shared housing market serves as an ideal context for investigating the role of social media information in discrimination: Firstly, approximately one-third of students in Germany live in shared apartments (Kroher et al., 2023). Secondly, it is rather common to share one's social media profile when searching a room (see Chapter 2; Moritz & Manger, 2022). Thirdly,

[17] Screenshots of the follower requests are presented in Figures B.12a (desktop view, p. 218) and B.12b (mobile app view, p. 218). These examples illustrate the requests from a profile for each ethnicity, although in the actual experiment, each user receives one request only.

reviewing such information is permissible under anti-discrimination laws (see also Chapter 2 of this thesis). Moreover, roommates are typically young adults and active users of social media platforms, making them inclined to consider such information during the selection process (de la Llama et al., 2012). Additionally, the subject pool closely aligns with that from Study I, enhancing the comparability of our findings.

The experiment was conducted between August and October 2023 in the 15 largest student cities in Germany using a 2x3 design. Each room advertiser is randomly assigned to receive an application[18] from either an ethnic minority or an ethnic majority name and one of three information conditions: In the first condition, only the application text is sent to an advertiser without any additional social media information (reference treatment). In the second condition, we include a corresponding social media profile without visual minority stereotypes, while in the third condition, we include a profile with visual minority stereotypes, e.g., images signaling religious affiliation and cultural identification.

The applicant text states that the fictitious applicant is 24 years old and studies business administration (the most common subject) in a master's program (Federal Statistical Office, 2022d). The application texts are the same across conditions. We add the most common hobbies that students mention on the platform, i.e., meeting friends, jogging, watching TV series, and traveling. In addition, we use characteristics that are often mentioned in application texts for shared apartments (Moritz & Manger, 2022), such as a reason for moving, non-smoker[19], and having experience in living in a shared flat.[20]

Furthermore, depending on the randomly assigned information condition, the application text includes a direct link to a social media profile with or without stereotypical pictures. This inclusion significantly reduces the cost and effort required for advertisers to collect additional information, making it immediately and easily accessible. In contrast to previous studies that implicitly assume employers or recruiters search online for applicants' information (Acquisti & Fong, 2020; Manant et al., 2019), our methodology embeds the social media profile directly within the application text – an important advantage of this rather informal, largely unregulated market (see also Chapter 2).

We apply to all ads that are on the platform for less than three hours at the time of the experiment, because advertisers in metropolitan areas with large student populations typically receive numerous applications within the first few hours of posting a vacant room. However, we only apply to rooms whose stated preferences and requirements – such as age and student status – match those of our fictitious applicant. We exclude rooms advertised by professional landlords, including real estate agents and rental agencies, because the selection decisions of professional landlords may differ from those of non-professional landlords, particularly in relation to the use of information from social media. Additionally, we did not apply to rooms available for short-term rentals (less than six months).

In addition, we implement a comprehensive duplicate check based on the information in the room ad, which filters out duplicate ads or advertisers who place multiple ads at the same time.

[18]The single application design mitigates the risk of detection.

[19]The social media profiles with visual stereotypes show pictures of the applicant smoking shisha. However, the text here refers to smoking at home, not leisure-time shisha smoking.

[20]A translation of the application text can be found in Section B.4 (p. 212).

This is to prevent advertisers from being treated twice and from realizing that they are part of an experiment. This should also minimize the costs to the subjects. We reject all responses within 24 to 48 hours, indicating that the fictitious applicant has already received an acceptance for another room on short notice.

To carry out the experiment, we write a computer program that collects available ads that meet the aforementioned criteria. The program then randomly selects the treatment conditions, generates the application and, after having been checked by one of the authors, sends it to the advertiser. The program also collects all the data about the characteristics of the room, the apartment, the advertiser, and the roommates, and automatically stores them in a database.

If the applicant receives a response, we collect these and classify them as a *callback* if the applicant is explicitly invited to a viewing with at least one of the potential roommates, or a (virtual) call. If an advertiser asks for more information, it is classified as *other* response. If the applicant is rejected for the room, it is classified as *rejection*.

We collect extensive data to use as control variables. This includes room and apartment characteristics such as size, monthly rent, and the number and sex of roommates (see Table B.16, p. 195). Additionally, we gather neighborhood and district-level socio-demographic variables, including the number of mosques and university buildings within a certain radius of the advertised room, as well as the proportion of foreign and Turkish residents in the respective district (see Table B.22, p. 202), and city-level election results – specifically, the share of votes for different parties and changes from the previous election – of municipal, state, federal, and European elections (see Table B.23, p. 203). We also collect data on the advertisers themselves, such as their sex and presumable ethnic origin. These contextual variables allow us to explore heterogeneity in unequal treatment.

Furthermore, we manually collect data on profile visits, impressions, and other engagement metrics for each profile throughout the experiment to verify whether the social media information is actually accessed (see Table B.24, p. 205).

3.4 Results

3.4.1 Study I: Formation of Informal Social Networks

3.4.1.1 Descriptive Results

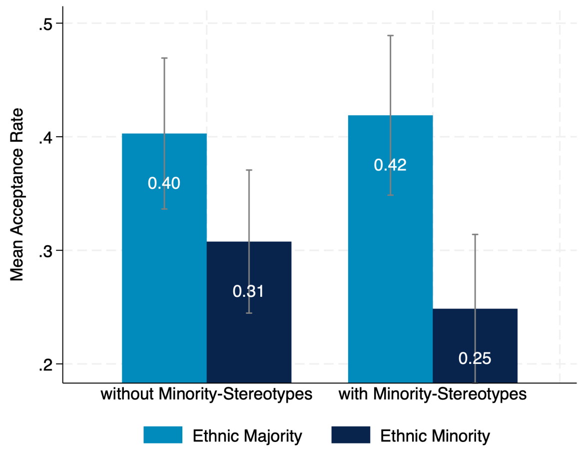


Figure 3.1: Mean Acceptance Rates

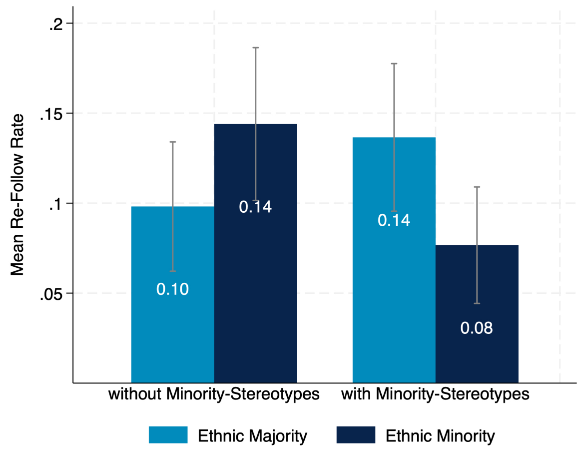


Figure 3.2: Mean Re-Follow Rates

When it comes to accepting friend requests on Instagram, Figure 3.1 shows a significant ethnic gap between minority and majority profiles that increases when profiles display visual minority stereotypes. While 31% of minority profiles without minority stereotypes and 22% of minority profiles with minority stereotypes are accepted as friends, majority profiles are accepted 40% and 42% of the time, respectively. This represents a statistically significant ethnic gap of 31% ($p = 0.042$) without minority stereotypes and 69% ($p = 0.001$) with minority stereotypes (see Table B.2, p. 178), indicating an increase in disparity when minority stereotypes are introduced.

Expressed in connections, ethnic minority profiles on average gain approximately 53 additional friends, while majority profiles gain around 83 additional friends. Notably, when minority profiles show minority stereotypes, their friend requests were accepted by an average of 22 fewer users compared to minority profiles without minority stereotypes. In contrast, majority profiles displaying minority stereotypes experienced a marginal decrease of only 5 fewer accepted friend requests on average (see Table B.2, p. 178).

In terms of re-follow rates[21], we find no significant difference between minority and majority profiles (see Figure 3.2 and Table B.2, p. 178). However, for profiles with minority stereotypes, the minority profile is less likely to be followed back, compared to majority profiles (14 vs. 8%) which constitutes a statistically significant difference ($p = 0.035$). In the absence of minority stereotypes, minority profiles are more likely to be followed back. However, this difference is not statistically significant (see Table B.2, p. 178). Interestingly, majority profiles are more likely to

[21] Private and public profiles can decide whether or not to re-follow the requesting profile. In our sample, 73.3% of suggested users have private profiles (see Table B.1, p. 177).

be followed-back when they present minority stereotypes (14 vs. 10%). However, this difference is again not statistically significant ($p = 0.1680$).

In terms of rejections, minority profiles are, on average, 48.2% more likely to be rejected compared to majority profiles ($p = 0.000$, see Table B.2, p. 178). When minority stereotypes are present, the rejection rate for minority profiles remains relatively stable at 28.7%, while majority profiles experience a decrease to 18.1%.

These findings underscore the substantial impact of visual stereotypes on inequality. The introduction of visual stereotypes exacerbates existing differences in acceptance and re-follow rates between ethnic majority and minority profiles.

3.4.1.2 Regression Analysis

We further investigate our results using different probit regression models to control for user characteristics that may affect the probability of being accepted or re-followed, respectively. Our main specification is as follows. For a fictitious profile i, the probability of acceptance (re-follow) is given by:

$$\begin{aligned} Pr(Accept_i(Re-Follow_i) = 1) = \Phi(\beta_0 &+ \beta_1 \times MajorityWithStereot_i \\ &+ \beta_2 \times MinorityWithoutStereot_i \\ &+ \beta_3 \times MinorityWithStereot_i + \gamma^{'}\mathbf{Z}_i), \end{aligned} \tag{3.1}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. $MajorityWithStereot_i$ is a treatment dummy variable which equals 1 if profile i is a majority profile with visual minority stereotypes, 0 otherwise. $MinorityWithoutStereot_i$ is a treatment dummy variable which equals 1 if profile i is a minority profile without minority stereotypes, 0 otherwise. Finally, $MinorityWithStereot_i$ is a treatment dummy variable which equals 1 if profile i is a minority profile with minority stereotypes, 0 otherwise. The reference category is the majority profile without minority stereotypes. $\mathbf{Z}_i$ is a vector of control variables for profile i (e.g., number of posts, followers, or account age). β_0 is the intercept, β_1-β_3 are coefficients for the main effects, and γ is a vector of coefficients corresponding to the control variables in $\mathbf{Z}_i$.

Table 3.1 presents the average marginal effects from probit regression models examining the determinants of both acceptance and re-follow rates. In the baseline model, without controls, sending a friend request from a minority profile with or without minority stereotypes has a large and statistically significant negative effect compared to a majority profile without minority stereotypes (-0.095, $p = 0.037$ without minority stereotypes and -0.159, $p = 0.001$ with minority stereotypes, see column 1 of Table 3.1).

When control variables are included, the negative effect for a minority profile with minority stereotypes remains significant, although the magnitude decreases. The effect for a minority profile without minority stereotypes becomes statistically insignificant, indicating that the inclusion of controls accounts for some of the variation observed in the baseline model. Interestingly, sending a request from a majority profile with minority stereotypes shows a positive and significant effect (0.086, $p = 0.067$, see see column 2 of Table 3.1), suggesting that minority stereotypes may slightly increase acceptance rates for ethnic majority profiles when user characteristics are controlled for.

Table 3.1: Determinants of Reaction Rates

	(1) Accept	(2) Accept	(3) Re-Follow	(4) Re-Follow
Ethnic Majority without Stereotypes (Ref.)	-	-	-	-
Ethnic Majority with Stereotypes	0.013	0.086*	0.037	0.059
	(0.046)	(0.047)	(0.027)	(0.041)
Ethnic Minority without Stereotypes	-0.095**	-0.010	0.043	0.192**
	(0.046)	(0.048)	(0.027)	(0.093)
Ethnic Minority with Stereotypes	-0.159***	-0.096*	-0.026	0.143
	(0.049)	(0.051)	(0.030)	(0.098)
Posts		0.000		-0.000
		(0.000)		(0.000)
Followers/subscribers		0.000**		0.000
		(0.000)		(0.000)
Following/subscriptions		0.000		0.000
		(0.000)		(0.000)
Mutual connections		-0.003		0.002
		(0.003)		(0.002)
Threads account		-0.001		0.050
		(0.064)		(0.032)
Account age		-0.001**		-0.001**
		(0.001)		(0.000)
Bio		0.073*		0.015
		(0.041)		(0.022)
Female		-0.111***		-0.064***
		(0.038)		(0.020)
Private		0.000		0.072**
		(0.000)		(0.030)
Turkish origin		0.361**		0.269***
		(0.165)		(0.066)
German origin		0.044		-0.013
		(0.039)		(0.021)
Observations	778	774	1,060	1,018
Pseudo R^2	0.016	0.079	0.011	-
Additional Controls	No	Yes	No	Yes

Note: The table reports average marginal effects computed from different probit models with accept as the dependent variable in columns 1 and 2 and re-follow in columns 3 and 4. Columns 1-3 report results from regular probit models, column 4 from a heteroscedastic probit model (see Section 3.4.1.4, p. 48, Table B.6, p. 184, as well as HECKMAN (1998) and NEUMARK (2012)). Columns 1 and 3 report the main effects (treatment conditions as dummies) without control variables. Columns 2 and 4 report the main effects including the displayed and additional control variables. Additional control variables that are not displayed include a dummy for whether the suggested user is new to Instagram, the number of username changes for a given user, the number of emojis in the bio, a dummy for having an active story, and the follower/subscriber ratio. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

With respect to user characteristics, sending a friend request to a female user has a significant negative effect on acceptance (-0.111, $p = 0.003$). However, an additional analysis of interactions with the dummy for female recipients (see Table B.8, p. 186) indicates that there is no statistically significant effect of the interaction between the treatment conditions and the user being female. Thus, while females might be more reluctant to accept requests from profile owners whom they do not know, the recipient's gender does not appear to influence unequal treatment in the formation of informal social networks.

A user with an alleged Turkish origin (as judged by his or her name) is significantly more likely to accept friend requests in general (0.361, $p = 0.029$). Moreover, when interacting Turkish origin and the treatment dummies, it can be seen that recipients with a Turkish origin are more likely to re-follow profiles displaying Turkish stereotypes, regardless of the sender's alleged ethnicity as signaled by the name. This implies a potential in-group preference or cultural affinity among Turkish-origin users when visual minority stereotypes are present (see Table B.7, p. 185).

In line with the descriptive results, the treatment conditions have no significant effect on re-follow rates in the baseline model without controls (see column 3 of Table 3.1). However, when control variables are included (see column 4), sending a friend request from a minority profile without minority stereotypes has a large and significant positive effect on being re-followed (0.192, $p = 0.039$) compared to the reference group. As before, the recipient being female negatively affects the likelihood of being re-followed, while a user having a Turkish origin positively affects the outcome.

Overall, the results provide causal evidence of inequality in informal social network formation, based on ethnicity and visual stereotypes. Ethnic minority profiles displaying minority stereotypes experience significantly lower acceptance rates, underscoring the prejudicial impact of visual stereotypes that signal cultural identification and religious affiliation. Conversely, when such stereotypes are absent, minority profiles engage more effectively, as seen in higher re-follow rates with controls included.

3.4.1.3 Profile Visits and Engagement

We analyze data on profile visits and other engagement metrics of the profiles to examine whether and how detailed profiles are viewed by potential friends under different experimental conditions. Figures B.1a (p. 180), B.1b (p. 181), and B.1c (p. 181) show box plots indicating the number of profile visits before, during, and after the experiment for the treatment conditions. Similarly, Figure B.2 (p. 182) plots the profile visits over time indicating the experimental periods. The figures show that during the experimental periods, when friend requests were sent, visits increased significantly, indicating that profiles are in fact accessed when receiving a friend request.

However, different treatment profiles receive different attention. Table B.4 (p. 180) presents summary statistics of weekly unique profile visits per request, segmented by ethnicity and visual stereotypes. Instagram only provides aggregate data on visits and other engagement variables. Therefore, we calculate average profile visits per request by dividing the number of visits to a given majority and minority profile in a given week by the number of requests for the respective name in the same week. Panel A of Table B.4 (p. 180) shows that majority profiles receive a higher average number of weekly profile visits compared to minority profiles. The difference is statistically significant at the 1% level. This indicates that potential friends are more likely to view the profiles of majority members than those of minorities.

Panel B of Table B.4 (p. 180) compares profiles with and without minority stereotypes. We do not find a significant difference in the average number of profile visits between these two groups. This finding is intuitively plausible, as users can only view the stereotypical images after accessing the profile, whereas the name indicating ethnicity is visible without having to visit the profile (see also Section 3.3.2, p. 40).

Regarding the interaction between ethnicity and minority stereotypes, Panel C indicates that majority profiles receive more profile visits compared to minority profiles. The difference is statistically significant at the 1% level. Similarly, Panel D shows that for profiles without minority stereotypes, majority profiles again receive significantly more visits.[22]

Furthermore, Figures B.3 (p. 193) and B.4b (p. 194) reveal that majority profiles tend to receive more impressions, i.e., posts are viewed more frequently, and have a higher reach of posts by non-subscribers[23]. Furthermore, profiles without minority stereotypes have a higher number of average daily impressions and reach as profiles with stereotypes.

These findings indicate that ethnicity significantly influences whether potential friends access and view profiles, with majority profiles mostly garnering more attention regardless of the presence of visual minority stereotypes. The lower number of profile visits to minority profiles imply that potential friends may be less inclined to explore profiles associated with minority groups. This reduced engagement could contribute to the observed inequality in informal social network formation, as fewer profile views may lead to fewer opportunities for interaction and connection for minority members.

3.4.1.4 Robustness Checks

Interaction of Ethnicity and Stereotypes The described effects in Section 3.4.1.2 (p. 45) remain robust, when including interaction effects of ethnic minority and minority stereotypes instead of treatment dummies (see Table B.3, p. 179). With acceptance as dependent variable (columns 1 and 2), the interaction term of a minority profile and minority stereotypes has a negative and statistically significant effect on the outcome when controls are included. This indicates that the combined effect of being an ethnic minority and displaying visual minority stereotypes significantly reduces the likelihood of a friend request being accepted, beyond their individual effects. Thus, the interaction drives the negative impact on acceptance rates, aligning with the previous findings that ethnic minority profiles with minority stereotypes face the most discrimination. Similarly, in the re-follow models (columns 3 and 4), the interaction term is consistently negative and statistically significant both without and with controls.

Heterogeneity of Friend Suggestions Regarding heterogeneity of friend suggestions, we test whether significant differences in user characteristics of potential friends suggested by the social media platform (see Table B.1, p. 177) affect the outcome by interacting these with the treatment conditions and regressing them on acceptance and re-follow rates. The interaction between the number of posts, a dummy for having a Threads account[24], and the length of the

[22] Comparing weekly and daily profile visits (see Table B.15, p. 193), we observe consistent patterns regarding ethnicity and the presence of stereotypes. However, for daily visits, minority profiles with stereotypes receive significantly more visits than ethnic majority profiles with stereotypes.

[23] Reach is defined as the number of unique users that have seen posts or stories at least once. It is different from impressions, which may include multiple views by the same user (see: https://help.instagram.com/788388387972460 and https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

[24] *Threads* is a social media platform which is closely linked to Instagram. It was introduced in 2023 by Meta Platforms, Inc. as an alternative to Twitter/X.

bio and all treatment conditions do not exhibit a significant effect on accept and re-follow (see Tables B.11, p. 189, and B.12, p. 190).

Similarly, interacting further characteristics from suggested friends (number of subscribers, mutual connections, new to Instagram, and account age) that show significant differences between minority and majority profiles with the treatment conditions, does not fundamentally change the main findings regarding the unequal treatment of minority and majority profiles (see Table B.13, p. 191, with accept and Table B.14, p. 192, with re-follow as dependent variable). While some interaction terms are significant, they primarily affect ethnic majority profiles, indicating that certain user characteristics may enhance or mitigate acceptance rates for this group. For ethnic minority profiles, the negative effect on acceptance rates mostly persists, underscoring the robustness of the discriminatory effect identified in the main analysis.[25]

Overall, the analysis demonstrates that heterogeneities in friend suggestions have negligible effects on unequal treatment. Some interaction terms reveal some nuances in how user characteristics influence acceptance rates. However, the (in most cases) persistent negative effect of ethnicity indicates that inequality is primarily driven by ethnicity and minority stereotypes rather than by differences in user characteristics.

Endogeneity Table B.5 (p. 183) presents a robustness check addressing a potential source of endogeneity that may arise if users accept friend requests because, in the meantime, a mutual friend has already accepted the same request. To control for this, the models include changes in mutual connections, capturing the number of new mutual connections between the time the friend request was sent and when the user reacted. The results indicate that the inclusion of changes in mutual connections does not affect the main outcomes. The interaction terms between changes in mutual connections and the treatment dummies are not statistically significant, indicating that the effect of the treatment conditions on acceptance and re-follow rates is not affected by changes in mutual connections. This confirms that the observed unequal treatment is not driven by users responding to new mutual connections but rather ethnicity and visual stereotypes.

Unobservable Heterogeneity Regarding unobservable heterogeneity, Table B.6 (p. 184) presents a robustness check addressing potential bias arising from unobservable heterogeneity between treatment groups. Specifically, we employ the Neumark correction by estimating heteroscedastic probit models to account for possible differences in error variances across groups, which could bias standard probit estimates (Heckman, 1998; Neumark, 2012). This approach helps to ensure that the observed effects are not driven by unmeasured factors that vary between ethnic minority and majority profiles or between profiles with and without minority stereotypes.

The results reveal that, after accounting for potential unobservable heterogeneity, the main findings remain largely robust. The significant positive effect of a minority profile without minority stereotypes on re-follow rates in the heteroscedastic probit model suggests that, when unobserved variance is considered, ethnic minority profiles without minority stereotypes may have a higher likelihood of being re-followed. This aligns with our earlier observation that the absence of minority stereotypes can mitigate unequal treatment.

[25] Additional (non-tabulated) results for the remaining significant differences (non-profit, active story, number of emojis in bio, and locations in bio) exhibit similar, non-significant results.

3.4.2 Study II: Inequality in the Housing Market

3.4.2.1 Descriptive Results

With respect to applications for shared housing ads, the descriptive results reveal a significant and consistent ethnic gap in favor of majority members/profiles. Figure 3.3 indicates that, without a link to a social media (SM) profile (reference treatment), the mean callback rate for ethnic minority applicants is 21%, compared to 35% for ethnic majority applicants, revealing an ethnic gap of 71%. This difference is statistically significant at the 1% level (see Table B.17, p. 197).

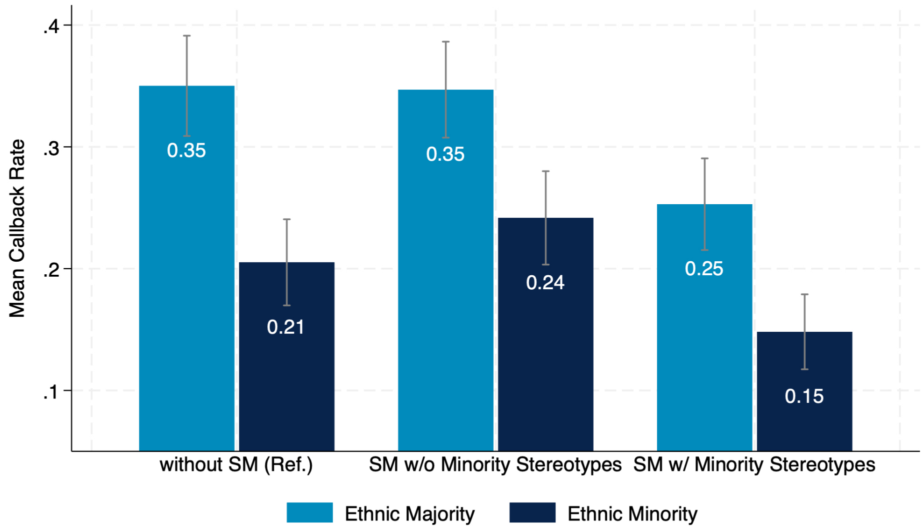


Figure 3.3: Mean Callback Rates

When applicants include a social media profile without minority stereotypes in their application, the ethnic gap decreases to 44% (statistically significant at the 1% level) with an average callback rate of 24% for minority and 35% for majority applicants. If applicants show minority stereotypes on their social media profile, the ethnic gap increases to 71% (again significant at the 1% level) with an average callback rate of 15% for minority and 25% for majority applicants (see Figure 3.3 and Table B.17, p. 197).

Analyzing the differences between the information treatments (see column 7 of Table B.18, p. 198), the results indicate that minority applicants need to send an average of 39% more applications to receive one callback when their social media profiles include minority stereotypes, compared to the reference treatment without a social media profile. Similarly, majority applicants need to send out 38% more applications when their profiles include minority stereotypes compared to the reference treatment. However, the negative effect of a social media profile that includes minority stereotypes is smaller for minority applicants than for majority applicants (see column 5 of Table B.18, p. 198).

These findings suggest that viewing profiles with visual minority stereotypes leads to a potential update in initial probabilistic beliefs, as majority applicants are not expected to conform to minority stereotypes, whereas minority applicants are more likely to be associated with such stereotypes. In other words, profiles with minority stereotypes affect how advertisers or roommates revise their initial assumptions about an applicant's characteristics (see Chapter 2).

For social media profiles without minority stereotypes, we find no significant difference from the reference treatment for majority applicants (see column 4 of Table B.18, p. 198). In contrast, minority applicants experience a positive effect, needing to send out 15% fewer applications to receive one callback compared to the reference treatment without a social media profile (see column 6 of Table B.18, p. 198).

This suggests that while majority profiles without minority stereotypes do not lead advertisers to update their initial probabilistic beliefs; for minority applicants, the absence of minority stereotypes is more likely to prompt an update of advertisers' (stereotyped) beliefs. Advertisers might actually expect to see stereotypical information associated with minority applicants. Therefore, profiles without minority stereotypes lead to a positive reassessment.

3.4.2.2 Regression Analysis

To further investigate our previous findings, we use various probit regression models that allow to control for different variables that may affect the probability of receiving a callback. Our main specification focuses on the interaction between ethnicity and the treatment conditions, which is as follows. For a fictitious applicant i, the probability of a callback is given by:

$$\begin{aligned} Pr(Callback_i = 1) = \Phi(\beta_0 &+ \beta_1 \times Minority_i + \beta_2 \times SM\_WithoutStereot_i \\ &+ \beta_3 \times SM\_WithStereot_i \\ &+ \beta_4 \times (Minority_i \times SM\_WithoutStereot_i) \\ &+ \beta_5 \times (Minority_i \times SM\_WithStereot_i) + \gamma' \mathbf{X}_i), \end{aligned} \tag{3.2}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution (probit link function). β_0 is the intercept. β_1-β_5 are coefficients to be estimated. $Callback_i$ is a binary variable that equals 1 if applicant i receives a callback, i.e, an invitation to a viewing and/or meeting with the potential roommate(s), 0 otherwise. $Minority_i$ is a dummy variable that equals 1 if applicant i has an ethnic minority name, 0 otherwise. $SM\_WithoutStereot_i$ is a binary indicator that equals 1 if applicant i includes a social media (SM) profile without minority stereotypes (Treatment 2), 0 otherwise, while $SM\_WithStereot_i$ is a dummy that equals 1 if applicant i includes a social media profile with minority stereotypes (Treatment 3), 0 otherwise.

$Minority_i \times SM\_WithoutStereot_i$ is the interaction term between minority status and a profile without minority stereotypes (Treatment 2). $Minority_i \times SM\_WithStereot_i$ is the interaction term between minority status and a profile with minority stereotypes (Treatment 3). In the reference category, applicants do not include a social media profile in their application (Treatment 1). $\mathbf{X}_i$ is a vector of control variables for applicant i (e.g., room and apartment attributes, roommate characteristics, advertiser information, geographic, and demographic controls, see Table B.16, p. 195, for summary statistics on the experimental variables). γ is a vector of coefficients corresponding to the control variables in $\mathbf{X}_i$.

Table 3.2 presents the average marginal effects from probit regression models estimating the probability of receiving a callback.[26] Across all specifications, having a minority name has a statistically and economically significant negative average effect on callback rates.

Including a social media profile without minority stereotypes (Treatment 2) does not significantly impact callback rates, suggesting that providing *neutral* additional information on average neither helps nor harms applicants. Including a social media profile with visual minority stereotypes (Treatment 3) significantly reduces the likelihood of receiving a callback for majority and minority applicants, but the negative effect is more pronounced for majority applicants – in line with the descriptive findings.

Table 3.2: Probit Models with Interaction Effects (Average Marginal Effects)

Callback	(1)	(2)	(3)	(4)
Ethnic Minority	-0.116***	-0.140***	-0.112***	-0.144***
	(0.0154)	(0.0235)	(0.0137)	(0.0179)
Treatment 1: Without SM (Ref.)	-	-	-	-
Treatment 2: SM without Minority Stereotypes	0.0152	-0.00436	0.0229	-0.00130
	(0.0236)	(0.0288)	(0.0252)	(0.0313)
Treatment 3: SM with Minority Stereotypes	-0.0797***	-0.0909***	-0.0564**	-0.0862***
	(0.0231)	(0.0287)	(0.0231)	(0.0271)
Ethnic Minority × Treatment 1 (Ref.)	-	-	-	-
Ethnic Minority × Treatment 2	-	0.0449	-	0.0447
		(0.0336)		(0.0292)
Ethnic Minority × Treatment 3	-	0.0253	-	0.0555**
		(0.0240)		(0.0220)
Observations	3,088	3,088	2,992	2,992
Month FE	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Room & Shared Apartment Controls	No	No	Yes	Yes
Roommate Controls	No	No	Yes	Yes
Advertiser Controls	No	No	Yes	Yes
Geographic & District Level Demographic Controls	No	No	Yes	Yes

Note: The table reports average marginal effects computed from different probit models with callback (invitation to a viewing) as the dependent variable. "SM" is the abbreviation for social media. Treatment 1 is the reference treatment and is therefore omitted. Columns 1 and 2 report results from regular probit models, columns 3 and 4 from heteroscedastic probit models (see Section 3.4.3, p. 55, Table B.27, p. 209, as well as HECKMAN (1998) and NEUMARK (2012)). Columns 2 and 4 report additional interaction effects of ethnic minority and treatment variables. Columns 3 and 4 report specifications with additional control variables. Monthly FEs include dummies for each month in which the experiment was conducted. City FEs include dummies for each city. Room & shared apartment controls include (among others) room size, total rent, number of roommates, online time, availability, and ad text characteristics. Roommate controls include roommate sex, languages spoken in the apartment, dummies for the shared apartment type, and other roommate characteristics. Advertiser controls include account age, a profile picture dummy, and advertiser origin (inferred from name). Geographic & district level demographic controls include (among others) distance to the city center, number of churches and mosques in the area, district population, share of female population, and Turkish population. Standard errors (in parentheses) are clustered at the city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The positive and significant interaction between minority and Treatment 3 (with visual stereotypes) in column 4 of Table 3.2 indicates that minority applicants suffer a smaller decrease

[26]For a detailed explanation of the callback classifications, please refer to Section B.3 (p. 211).

in callback rates when including stereotypical images compared to majority applicants. In other words, the marginal penalty of displaying visual minority stereotypes is smaller for minority applicants, possibly because advertisers may have adjusted their beliefs regarding stereotypical associations with minority groups as described in the previous section.

The results are robust to the inclusion of a large number of control variables on district, room, and apartment characteristics, as well as monthly and city dummies. Furthermore, the effects are robust to using split samples with respect to ethnicity (see Table B.19, p. 198). The split sample results are consistent with those from the main model incorporating interaction terms. Both analyses indicate that including a social media profile with minority stereotypes reduces callback rates more substantially for ethnic majority applicants than for minority applicants.

Regarding rejections, Table B.26 (p. 206) presents the results of different probit regressions estimating the probability of receiving a rejection. Across all specifications, the coefficients for minority applicants are positive but small and not statistically significant. This indicates that applicants with minority names are not significantly more likely to receive rejections compared to ethnic majority applicants. Similarly, the coefficients for Treatment 2 and 3 are small and lack statistical significance in all models. The interaction terms are likewise insignificant. These findings suggest that neither the applicant's ethnicity nor the inclusion of social media profiles – with or without minority stereotypes – significantly affect the likelihood of receiving a rejection.

We further analyze potential determinants of inequality by using a range of contextual factors, including room, apartment, roommate, and advertiser characteristics, as well as geographic and demographic variables. The split sample analysis in Table B.19 (p. 198) suggests that certain city-level features and cultural signals may moderate discrimination for minority applicants. For instance, non-tabulated results indicate that a larger local Turkish population is associated with a significantly positive effect on callback rates for minority applicants. However, using an interaction term of minority and the share of Turkish population in the district reveals that while a higher Turkish population share generally correlates with more callbacks, Turkish applicants themselves do not benefit equally. In other words, areas with a larger minority community might be more welcoming overall, but the relative advantage this offers is actually smaller for minority applicants than for majority applicants.

The advertiser's origin also appears to matter. Advertisers who – given their name – allegedly originate from a Muslim-majority country are associated with a significantly higher callback rate for minority applicants, suggesting that cultural similarities may facilitate market access for minority applicants. Yet, when using interaction terms, this effect is still positive, but no longer statistically significant.

Interestingly, a higher number of female roommates has a significant negative effect on callback rates, which is slightly more pronounced for minority applicants. Nonetheless, the interaction between minority status and the number of female roommates does not show statistical significance, suggesting that the roommates' gender does not drive unequal treatment.

Finally, an interaction term between minority status and the number of nearby mosques shows a positive and significant effect, partially offsetting the generally negative effect of mosques on callback rates. This suggests that while having more mosques nearby is associated with fewer callbacks in general, for minority applicants this negative association is attenuated.

Overall, the results demonstrate that both ethnicity and the (absence of) visual minority stereotypes significantly affect selection decisions. Minority applicants are consistently less likely to receive callbacks, highlighting systemic unequal treatment. While providing a social media profile without minority stereotypes does not significantly affect callback rates on average, minority stereotypes exacerbate the negative effects for majority applicants but less so for minority applicants. Most contextual factors that might be drivers of unequal treatment are either not significant or not robust, suggesting that the observed unequal treatment is more fundamentally rooted in ethnic affiliation and the (absence of) minority stereotypes than in variations of local demographic conditions, roommate, or advertiser characteristics.

3.4.2.3 Profile Visits and Engagement

As part of Study II, we also collect weekly profile statistics on followers, likes, visits, impressions, and many others. Because 91.7% of the responses are received within one week of the application, weekly visits should provide an accurate estimate of profile visits per application.

As in Study I, Instagram only provides aggregate engagement data, so we cannot infer the exact proportion of advertisers who actually clicked on the profile link in the application.[27] Therefore, we again calculate average profile visits per application by dividing the number of visits to a given majority and minority profile in a given week by the number of applications containing a social media profile for the respective name in the same week.

The results presented in Table B.20 (p. 199) indicate that, on average, majority profiles receive significantly more profile visits (0.19 vs. 0.13) and impressions (1.47 vs. 0.79) per application, indicating that ethnic majority applicants garner greater engagement and attention from potential roommates. The differences are statistically significant and substantial in magnitude, with ratios of around 1.48 for profile visits and 1.85 for impressions, suggesting that ethnic majority profiles benefit from a pronounced engagement advantage. Likewise, the reach of non-subscribers is approximately 58% higher for ethnic majority profiles compared to minority profiles.[28]

These findings highlight that social media information is indeed accessed and utilized by advertisers, and that advertisers' engagement is shaped by applicants perceived ethnicity. By receiving fewer visits and impressions, ethnic minority applicants may face a double disadvantage: not only are they less likely to receive callbacks, but they also see lower levels of profile engagement that could otherwise help them stand out or signal suitability. Consequently, the role of profile visits and other engagement variables is not simply a neutral process of additional information gathering, but one that can perpetuate inequality and even exacerbate existing biases.

[27] Linking each individual advertiser to the number of visits would be technically feasible if the link to the social media account contained a unique ID, as in Bartoš et al. (2016) who use personal websites. However, we consider the risk of detection to be extremely high, as it would not be technically possible to hide the ID in the link. Furthermore, this would indicate that clicking on the link would first open another website that tracks the click. This could make advertisers suspicious.

[28] See also Figures B.5a (p. 200) and B.5b (p. 200) showing profile visits per treatment condition during and outside the experiment.

To investigate the drivers of weekly profile visits, reach, and impressions, we regress visits, reach, and impressions on selected roommate and apartment characteristics and social media variables using OLS models (see Table B.21, p. 201). The results reveal a significant negative effect of a minority name on all engagement metrics. In other words, ethnic minority applicants see substantially fewer profile visits, lower reach among non-subscribers, and fewer impressions, indicating that minority applicants are at a considerable disadvantage in terms of online engagement.

The interaction of minority status and Treatment 3 (social media profile with visual minority stereotypes) is positive and large in magnitude, more than offsetting the individual negative effects. This suggests that while minority profiles generally face reduced engagement, minority stereotypes actually increase engagement. One possible interpretation is that advertisers or potential roommates might visit and view a minority profile with minority stereotypes more often and in more detail. Furthermore, they might more often pass on the profile link to other roommates. Additional variables reveal that having a temporary room listing reduces engagement while ads that explicitly state that a social media profile should be included in the application, increases engagement significantly.

3.4.3 Robustness Checks & Additional Analyses

Unobservable Heterogeneity The additional robustness check using heteroscedastic probit models, shown in Table B.27 (p. 209), addresses potential bias arising from unobservable heterogeneity between treatment groups. By employing the Neumark correction, which involves estimating heteroscedastic probit models to account for differences in error variances across groups (Heckman, 1998; Neumark, 2012), we ensure that the observed effects are not driven by unmeasured factors varying systematically between minority and majority applicants or between profiles with and without minority stereotypes.

The associated Wald tests for equal variances do not indicate a variance difference between minority and majority applicants in columns 1 and 2. However, the model in column 3 yields a significant p-value, suggesting that the variance of unobserved characteristics differs once a more comprehensive set of controls is included, especially room & apartment and advertiser controls. Nonetheless, even after accounting for these differences in unobservables, the direction and significance of the treatment and interaction effects remain robust. However, to address differences in unobservables, we estimate heteroscedastic probit models instead of regular probit models when necessary, see for example column 4 of the main model (see Table 3.2).

Voting Behavior & Discrimination Table B.28 (p. 210) explores a possible mechanism of ethnic discrimination, namely, whether regional voting patterns for right-wing parties affect unequal treatment of minority applicants. The probit models include interaction terms between the minority status of applicants and the share of votes for right-wing parties, such as the "AFD" ("Alternative für Deutschland") and "NPD" ("National Democratic Party of Germany")[29], from various elections (municipal, state, federal, and European). The dependent variable is the likelihood of receiving a callback.

[29]The party renamed itself in 2023 from "NPD" to "Die Heimat".

The results indicate that minority applicants consistently receive fewer callbacks across all models. Notably, in the state elections (column 3), the interaction between ethnic minority status and AFD vote share is negative and statistically significant ($-0.00326, p = 0.091$), suggesting that in regions with higher support for the AFD, minority applicants face a larger reduction in callback rates compared to majority applicants.

Similarly, in the federal elections (column 6), the interaction between minority status and NPD vote share is negative and significant ($-0.548, p = 0.009$), indicating that higher NPD support correlates with increased discrimination against minority applicants. These findings imply that regional political climates, as reflected by voting results for right-wing parties, may exacerbate unequal treatment in the housing market, leading to fewer opportunities for ethnic minority individuals.

3.5 Limitations

Our research design has limitations. One limitation that affects the generalizability of our findings relating to the methodology used to create the fictitious social media profiles. We designed these profiles to contradict or conform to prevailing ethnic minority stereotypes, as held by the ethnic majority. However, the artificial construction of applicant profiles, as well as the controlled variation in visual stereotypes, may not perfectly capture the complexity of real-life situations. Actual applicants could present a more dynamic array of signals, such as subtle linguistic cues, additional credentials, or longer online histories, which might be interpreted differently by suggested friends or potential roommates.

Another limitation concerns the inability to observe the underlying mechanisms driving the observed behavior. While our findings suggest that suggested friends or potential roommates rely on ethnic signals and social media content to form probabilistic beliefs, the research cannot directly measure the decision-making processes, motives, or implicit biases at play. The data only capture outcomes (acceptance, re-follows, callbacks, profile visits, etc.) rather than the full cognitive processes behind them. Additionally, although various robustness checks and the Neumark correction help mitigate issues arising from unobservable heterogeneity, there may still be unmeasured factors influencing both the formation of informal networks and potential roommates' selection criteria.

While we interpret our findings on unequal treatment as predominantly indicative of belief-driven statistical discrimination, with social media information appearing to challenge initial beliefs about a minority applicant's undisclosed characteristics, we are unable to discern the specific contributions of Bayesian statistical discrimination, non-Bayesian statistical discrimination, or unexplained/taste-based discrimination to the observed ethnic gap. Moreover, the preconceived stereotypical perceptions of suggested friends or potential roommates may be so deeply ingrained, possibly due to persistent inaccurate statistical assumptions about group identity (Bohren et al., 2019; Campos-Mercade & Mengel, 2023), that additional social media information may play a comparatively minor role.

Moreover, to limit treatment arms and since stereotypes may vary in extent and frequency by gender, we concentrated on male applicants only. However, future research might also incorporate a variation in gender to investigate the combined role of visual stereotypes, ethnicity, and gender.

3.6 Conclusion

Previous evidence suggests that the provision of additional information significantly affects market outcomes of candidates applying for jobs, housing, or other (Bertrand & Duflo, 2017; Cui et al., 2020; Hermes et al., 2024; Moritz et al., 2023; Nunley et al., 2011). However, there is no experimental evidence for the role of stereotypes as additional information in economic or social settings, particularly regarding visual representations of commonly held minority stereotypes. This is despite the growing relevance of information presented on social media platforms over the past decade (Acquisti et al., 2015), as evidenced by employers increasingly utilizing these platforms to identify prospective employees (Becton et al., 2019; Rosen et al., 2018) and assess candidates' suitability for hiring decisions (Roth et al., 2016).

We conducted two field experiments to investigate the combined effect of ethnicity and visual stereotypes on unequal treatment in the formation of informal social networks and on callback rates in the market for shared apartments. The results of both studies – using a demographically very similar subject pool – indicate significant ethnic gaps in the acceptance rates of friend requests and callback rates for vacant rooms. We find that minority members are significantly less likely to be accepted as digital friends (46% gap) and less likely to be called back as potential roommates (61% gap).

More importantly, visual narratives significantly increase inequality, especially in the formation of informal social networks (69% gap compared to 71% gap in callback rates for vacant rooms). However, conforming to minority stereotypes does not affect majority members regarding friend requests (40 vs. 42%), while it significantly affects majority members regarding callbacks (35% vs. 25%).

In shared housing, the marginal penalty of displaying minority stereotypes is smaller for minority applicants, potentially due to advertiser's adjusted expectations with respect to stereotypical associations with minority groups. This indicates that while majority profiles without minority stereotypes do not prompt advertisers to revise their initial probabilistic beliefs, minority applicants are more likely to trigger an update of advertisers' (stereotyped) beliefs when their profiles lack minority stereotypes, contradicting common beliefs by the majority.

While previous literature on the formation of professional (job) networks (Ajzenman et al., 2025; Evsyukova et al., 2025) reported racial gaps of 12% and 13%, respectively, our results indicate an ethnic gap of 31% for profiles without minority stereotypes and even 69% with minority stereotypes in the formation of informal social networks. This suggests that informal social interactions may be more susceptible to unequal treatment and stereotypical perceptions than those in rather formal, professional settings.

Additionally, in both studies, the inequality in treatment between minority and majority members is lowest for applicants who do not show minority stereotypes on their social media profiles. However, in contrast to the findings in Chapter 2, the ethnic gap is not closed in this case.

Both studies show that social media information is actively accessed and used for decision-making with significant effects of ethnicity on profile visits and engagement. Majority profiles consistently attract more attention regardless of whether or not they display visual minority stereotypes. In addition, average weekly visits are much smaller for potential roommates than for potential friends. There is a consistent and significant ethnic gap in visits, although in both studies this gap is smaller when minority stereotypes are shown.

Nevertheless, the reduced engagement with (stereotypical) minority profiles suggests that potential friends and roommates are less inclined to explore minority profiles, limiting opportunities for interaction and connection – which reduces their chances of signaling suitability. Consequently, visiting a social media profile and examining its contents is not a neutral process of additional information gathering; it rather can perpetuate inequality and exacerbate existing biases.

Because the costs associated with accepting a friend online are significantly lower than those involved in meeting a potential roommate in one's home, it may be economically rational to discriminate less in online social networks. Still, as it takes only a few seconds to dissolve a friend connection online, discrimination rates are proportionally high in the online setting of Study I. One reason for this may be that the composition of an individual's online social network is a signal in itself and may be perceived as a costly signal regarding impression formation.

Overall, both studies demonstrate the critical need for targeted interventions to address and mitigate implicit biases across both digital and real-world outcomes. While further research is needed to determine how best to implement such measures, local housing authorities, anti-discrimination legislators, and platform regulators could all play a role in addressing subtle forms of bias. For instance, platforms could consider incorporating features designed to neutralize group-based signals, while legislators could expand anti-discrimination legislation to encompass informal settings while ensuring its enforceability. By acknowledging the particular vulnerability of ethnic minority members in informal social and economic settings, these stakeholders can develop more effective strategies to promote inclusive environments across all areas of social and economic life.

Chapter 4

The Effects of Information Salience on Ethnic Discrimination[1]

4.1 Introduction

The salience of information shapes perception and affects human behavior in many ways (Bordalo et al., 2013, 2022; DellaVigna, 2009; Tversky & Kahneman, 1974). Think of an eye-catching display in front of the cheese counter in a supermarket showing handmade French Camembert with a glittering golden seal indicating a high quality product from the green and gentle hills of Normandy – you now might consider to buy (instead of a cheaper one). Or a flashing button on an online travel platform indicating that hundreds of visitors are about to book the same hotel as you, so you'd better hurry.

Research indicates that information salience affects decision-making processes, as more salient attributes receive disproportionate weight in human behavior regarding consumer choices (Bordalo et al., 2012, 2022; Chetty et al., 2009; Michels et al., 2024), investment decisions (Frydman & Wang, 2020), and political beliefs (Bordalo et al., 2020). In the growing world of online marketplaces and matching platforms, where prospective employees, tenants, or roommates are often listed in an order determined by platform design (and sometimes by pay-to-feature schemes), understanding the role of visibility and salience is crucial as information salience defines the way individuals form judgments of others (Hamilton et al., 1978) or discriminate against ethnic minorities (Doleac & Stein, 2013; Manant et al., 2019).

In this paper, I present evidence from a field experiment designed to investigate the role of information salience in affecting ethnic discrimination in the housing market. A vast literature in economics and related fields has documented pervasive ethnic discrimination in housing, labor, and other markets. Classic correspondence studies (Ahmed et al., 2010; Bertrand & Duflo, 2017; Bertrand & Mullainathan, 2004; Lippens et al., 2023) reveal that applications bearing distinctively minority names receive systematically fewer callbacks than those with majority

[1] A very similar version of this field experiment (see Chapter 3) was approved by the Ethics Committee of the Faculty of Economics and Social Sciences of the University of Tübingen (A2.5.4-286_bi) on May 2, 2023. No additional IRB approval was obtained for the additional treatment condition. The research project was pre-registered in the AEA RCT Registry (#11322 and #14294): https://doi.org/10.1257/rct.11322-3.0 and https://doi.org/10.1257/rct.14294-1.0.

names, all else being equal. While previous research has extensively examined ethnic discrimination in various settings, relatively little attention has been paid to how information salience – that is, the visibility or prominence of an application (think of the golden seal from above) – shapes the extent of unequal treatment.

To investigate the effect of information salience on unequal treatment, I adopt the experimental design described in Chapter 3 and send 4,270 fictitious applications to vacant room ads on an online matching platform for roommates in the largest German student cities. I randomly vary the subscription status of the applicant, i.e., whether applications are sent from paid "premium" accounts or from non-premium accounts on the roommate platform.

Applications from premium accounts are visually highlighted by the platform and gain a significantly higher position in the advertiser's or roommate's inbox and thus more salience. Premium applications are placed near the top of the queue (on average at position 3), whereas non-premium applications appear much lower (on average around position 16). In this way, this unique platform setup allows me to interpret inbox position as a proxy for information salience: Top-listed applications are more likely to capture an advertiser's attention, while those that appear farther down the list are at risk of being skimmed or overlooked, and can be expected to receive fewer callbacks. In addition, the experiment is split in two distinct waves: the first occurs shortly after the launch of the platform's new feature, while the second takes place one year later. Examining the differences between these two periods enables an investigation of potential first-mover effects, as well as long-term adaptations in user behavior to this platform's monetization strategy. Furthermore, I randomly vary an applicants' ethnicity – signaled by names associated with minority or majority group members – and further manipulating whether the application contains a link to a social media profile and whether this profile exhibit content related or unrelated to minority stereotypes (see Chapter 3). Thus, my study presents causal evidence on how high vs. low information salience (i.e., premium vs. non-premium placement) affect callback rates in general and ethnic differences in callbacks in particular.

To the best of my knowledge, the present study is the first to investigate how varying the salience of an application affects market outcomes for ethnic minority and majority members. While numerous correspondence studies in labor and housing markets have documented discrimination by manipulating applicant characteristics (e.g., Acquisti & Fong, 2020; Bartoš et al., 2016; Bertrand & Mullainathan, 2004), they typically have limited control over the visibility of applications, making it nearly impossible to isolate the effect of an application's salience. Moreover, I contribute to the growing body of field experiments exploring how digital platform features interact with discrimination (Doleac & Stein, 2013; Edelman et al., 2017).

The closest to my research is a study by Manant et al. (2019) that examines ethnic discrimination in hiring. In this study, an applicant's origin is revealed solely on a social media profile. Manant et al. (2019) find an ethnic gap of approximately 42%, suggesting that recruiters are indeed searching online for supplementary information about applicants. During the course of the experiment, the social media platform modified the profile layouts, reducing the salience of origin information. Subsequently, the ethnic gap disappeared. However, their approach does not allow for random variation of a premium status that affects inbox position, nor does it permit integrating links to the respective social media profiles directly into the application itself. By contrast, the unique setup of my experiment enables explicit, randomized manipulation of both salience and the portrayal of minority identities, allowing to investigate whether – and to what extent – salience and attention-based mechanisms might reinforce, diminish, or fail to affect discriminatory behavior.

My results indicate that premium status, i.e., higher salience of an application, increases the average callback rate by about 3 percentage points (significant at the five percent level). Non-premium applications with lower salience need to send on average 11% more applications to receive one callback compared to premium applications with higher salience. I do not find evidence for any kind of first-mover advantage in upgrading to a premium status shortly after the introduction of the feature.

Interestingly, across all information treatments, higher salience does not significantly affect the ethnic gap between minority and majority applicants, indicating that moving closer to the top of the inbox does not significantly affect discriminatory behavior. While both, ethnic minority and majority applicants gain a slight increase in callback rates from higher information salience, premium status raises callbacks by approximately 26% for minority applicants with a social media profile free of minority stereotypes and by 18% for majority applicants in the same condition. The differences are marginally statistically significant at the 10 percent level. In contrast, minority applicants who post minority stereotypes on their social media profile do not gain from having a premium account: their callback rate is about one percentage point lower than the one of their non-premium counterparts.

Moreover, the findings indicate that each single-position drop in the inbox is associated with an approximately 0.3 percentage-point decline in the callback rate. Although this may appear modest, it accumulates substantially over multiple positions. While prior results do not reveal a robust overall effect of salience on discrimination, they do show that lower inbox positions impose a disproportionately larger penalty on minority applicants – especially within the part of the inbox where applications are still likely to be read rather than dismissed.

Overall, the findings indicate that while having a premium account to increase visibility can slightly increase callbacks, it neither eliminates the persistent disadvantage faced by minority applicants nor meaningfully weakens the relationship between ethnicity and callback rates. The results thus point to salience as a mechanism that amplifies attention for all applicants but leaves underlying discriminatory biases largely intact.

The findings bear important implications for policy and platform design (Cheng et al., 2024). Making an application from an ethnic minority member more salient is not helpful in ensuring equitable opportunities. Such results speak to how private platforms might inadvertently shape minority outcomes in housing or labor markets, and suggest potential interventions – like regulating pay-to-feature options or designing platform features or algorithms that are less bias-susceptible (Lambrecht & Tucker, 2019).

The remainder of this chapter is organized as follows. Section 4.2 provides a theoretical framework outlining how salience alone – and in combination with ethnicity – may affect callback rates. Section 4.3 describes the experimental design. Section 4.4 presents the empirical strategy and main results, while Section 4.5 describes limitations of the present research. Section 4.6 discusses the results and concludes.

4.2 Theoretical Framework

4.2.1 Information Salience

The concept of salience (Bordalo et al., 2012, 2022; Chetty et al., 2009) posits that individuals disproportionately attend to attributes or signals that are presented prominently (e.g., top search results, or front-page listings). In online matching environments, information salience shapes how decision-makers sort through potential matches. When many applicants compete for attention, the position at which an application appears – and the degree to which certain signals (e.g., profile attributes) are highlighted – can profoundly affect success of matching.

In the context of an online (shared-)apartment platform, premium applications are placed at a higher position in the advertiser's, roommate's, or landlord's inbox, thereby increasing their probability of being opened and read. The theoretical mechanism rests on attention constraints: advertisers often examine only the first few applications in detail, either due to time scarcity or cognitive overload (Bartoš et al., 2016; DellaVigna, 2009; Simon, 1955). Applications that appear "below the fold" may be skimmed or overlooked. Consequently, premium (higher) positions enjoy better visibility and are more likely to generate callbacks. The concept of salience offers a useful framework to investigate how attention is captured in such scenarios. Specifically, Bordalo et al. (2022) describe salience as "the property of a stimulus that draws attention bottom up" (Bordalo et al., 2022, p. 524), indicating three prerequisites that make a stimulus salient: *contrast*, *surprise*, and *prominence*. I argue that sending a "premium" application which lead to a top inbox position renders it highly salient along these three key dimensions.

Firstly, as premium applications not only occupy a conspicuously high position, while non-premium applications appear much lower, they are also conspicuously highlighted. This creates a visual *contrast* in the advertiser's inbox: the top set stands out against lower-ranking applications. Since human attention is drawn to contrasts (Bordalo et al., 2022), having a block of premium listings distinctly separated from the larger volume of non-premium listings makes these top-positioned applications more salient.

Secondly, in many online marketplaces, one typically sees a chronological or alphabetical list of incoming messages. By contrast, a premium application leaps ahead of the queue, which is likely to violate the advertiser's expectation for how messages are ordered. The mismatch between the typical (date-based) ordering and an artificially boosted (premium) ordering can generate a *surprise*, nudging the advertiser's attention towards out-of-place items.

Thirdly, salience theory posits that highly prominent attributes (or items) get over-weighted in the decision process (Taylor & Thompson, 1982). According to Bordalo et al. (2022), an option's *prominence* (e.g., being displayed first and flagged with special icons) increases the chance of bottom-up attention allocation. As premium messages appear first on the screen, effectively taking the spotlight in the advertiser's field of view, advertisers are thus more likely to read it carefully, recall its details, and respond to it. By contrast, non-premium applications become less salient and more prone to "distraction", i.e., they risk being under-weighted or even ignored.

Prediction 1. *Hence, I expect that premium (top-position) applications will generally enjoy higher callback rates than non-premium (lower-position) applications, driven by enhanced visibility, lower "reading costs," and increased attentional focus. Specifically, the more salient the application, the more likely advertisers are to open, read, and respond to it.*

By systematically comparing callback rates for premium and non-premium applications, I can test whether salience alone, independent of any differences in applicant characteristics, meaningfully boosts the prospects of receiving a callback. This prediction aligns with empirical findings on attention constraints (Bartoš et al., 2016; DellaVigna, 2009) and extends them by examining a subscription-based mechanism that directly affects an application's rank and prominence.

4.2.2 Information Salience & Ethnic Discrimination

A large body of research in labor and housing markets demonstrate that minority applicants often face systematic discrimination, reflected in lower callback rates relative to majority applicants (Bertrand & Duflo, 2017; Lippens et al., 2023). In classical models of taste-based discrimination (G. S. Becker, 1957), majority-group employers (advertisers) derive disutility from interacting with minority applicants and thus are less likely to respond. By contrast, statistical discrimination theories (Arrow, 1973; Phelps, 1972) posit that employers (advertisers) use group membership as a proxy for unobservable characteristics (e.g., reliability, sociability). Both frameworks predict unequal outcomes when ethnicity is salient.

Information salience can either attenuate or exacerbate ethnic differences in callbacks. On the one hand, salience might help minority applicants overcome negative priors if they appear first, catching the advertiser's attention before biases fully manifest. On the other hand, if ethnic signals remain prominent (for instance, via names or social media information), salience might merely accelerate discriminatory behavior – an advertiser who disfavored minority applicants could quickly reject them, even at a higher position. Which outcome prevails also depends on how the marginal increase in attention interacts with advertisers' preferences or beliefs about minority applicants.

An additional element is the information treatment embedded in social media profiles. Advertisers, potential roommates, landlords, or employers might search for additional information about applicants online (Acquisti & Fong, 2020; Manant et al., 2019) or this information is directly accessible within the application (see Chapters 2 and 3 of this thesis). Profiles can be absent, neutral, or stereotype-laden in their portrayal of minority status, thus manipulating the salience of ethnic signals. In classical frameworks, when more precise group signals are available, statistical discrimination intensifies if advertisers believe these signals confirm a negative stereotype. If, however, the additional information is neutral or contradicts common stereotypes, it could mitigate negative biases (see Chapter 2 of this thesis). Meanwhile, under taste-based discrimination, providing more information need not change attitudes, as a preference-based distaste remains present regardless of new information (G. S. Becker, 1957).

Differences in information salience might affect minority applicants in different ways. If the additional visibility allows them to demonstrate attractive qualities or to challenge stereotypes (e.g., a personal social media profile free of negative cues), salience may reduce discrimination (Benjamin et al., 2010; Bertrand & Duflo, 2017). Conversely, if taste-based or statistical biases are strong and the applicant shows stereotypical information (e.g., a personal social media profile including stereotypical images), a more prominent minority and stereotypical label can trigger a quicker negative decision, worsening differential treatment.

Social media profiles introduce another dimension of information – they can be perceived as positive/contradictory to stereotypes or negative/stereotypical while the absence of such information could be perceived as neutral. Under statistical discrimination (Arrow, 1973; Phelps, 1972), advertisers use observable traits to update beliefs about the applicant's overall "quality" (Altonji & Pierret, 2001). If the profile contains negative cues, it could confirm prior biases, heightening discrimination. If it is neutral or positive, it might offset negative group priors, particularly if attention is directed to those details early in the decision process (H. Fang & Moro, 2011). Under taste-based discrimination (G. S. Becker, 1957), additional information may not shift prejudiced preferences at all, but it could still affect the speed or likelihood of a callback if the advertiser invests more or less effort in reading applications from minority members.

In sum, the salience of an application – operationalized by inbox position – can interact strongly with both ethnic signals and profile content. When minority cues are subtle, a higher inbox position might lead to fairer evaluation. When stereotypes are salient, a higher inbox position could magnify negative bias. This dual mechanism aligns with existing theoretical and empirical work on taste-based and statistical discrimination (e.g., G. S. Becker, 1957; Phelps, 1972), which suggests that beliefs or prejudices about a minority group's productivity or desirability can be activated more strongly when relevant signals are made conspicuous.

Prediction 2. *Hence, I expect that the marginal benefit of a top position (see Prediction 1) may vary systematically by ethnicity and information treatments: (a) minority applicants are likely to gain less from higher salience (higher placements) compared to majority applicants as ethnic signals are more prominent and therefore, may trigger stereotypical beliefs. However, (b) minority applicants whose profiles lack stereotypical cues are expected to gain a greater boost from high-salience placements, since advertisers/potential roommates are more likely to read and update prior beliefs favorably, but (c) minority applicants who inadvertently highlight stereotype-consistent or otherwise "negative" characteristics may see this advantage evaporate – or even reverse – if the increased attention reinforces detrimental assumptions.*

The present chapter describes a randomized field experiment investigating how the intersecting factors – salience, ethnicity, and additional (non-)stereotypical social media information – jointly affect callback decisions. I examine the relationship between cognitive attention constraints and biases in an online environment where decision-makers face cognitive and time constraints that limit how thoroughly they evaluate each applicant as reviewing every application in detail is costly, advertisers often focus attention on the first few (highly salient) applications presented (Bartoš et al., 2016; Gabaix, 2014; Simon, 1955).

4.3 Experimental Design

To empirically test the predictions, I conduct an experiment on Germany's largest online platform for shared apartment ads, which functions similarly to Airbnb but focuses primarily on long-term rentals. In August 2023, the platform introduced a subscription model, allowing users to upgrade to a premium account by paying either 13.90 Euros per month for a 12-month subscription or 20.90 Euros for a single month. Applications from premium users are visually highlighted and appear at the top of the advertiser's inbox.[2] Furthermore, premium users gain access to advertisement metrics, such as the number of users who have viewed the ad and the total number of (non-)premium applications sent.

Using this data, I observe actual inbox positions for premium and non-premium applications, offering an innovative top-versus-bottom manipulation to examine how high versus low salience affects "success," i.e., whether an applicant receives an invitation to view the room and/or meet the potential roommate(s), as measured by callback rates. Since premium applications are both visually highlighted and occupy significantly higher inbox positions, this design effectively captures bottom-up attention (Bordalo et al., 2022).

Although many online apartment and room rental platforms require landlords (supply side) to subscribe to a premium service or pay a one-time fee to list their property, premium subscriptions for potential tenants (demand side) – allowing them to contact advertisers, enhance their visibility, or use other premium features – are less frequent.[3] Moreover, only a few platforms feature a higher salience of an application to paying subscribers, making this market and the platform particularly well suited for investigating the effect of information salience on market outcomes.

I study the effects of salience, ethnicity, and (non-)stereotypical social media information by randomly assigning applications to premium or non-premium accounts[4]. The remaining experimental design is based on Chapter 3 of this thesis, using social media profiles on the social media platform Instagram depicting a 24-year-old male master's student living in Germany. Ethnicity is signaled through names: half of the applications and respective social media profiles bear distinctively Turkish-sounding (minority) names, and the other half German-sounding (majority) names.[5] Additionally, to either contradict or conform with common Turkish stereotypes, half of these profiles feature stereotypical images (e.g., visits to a mosque, smoking shisha with (male) friends, posing in a fast car, displaying the Turkish flag). These "visual narratives" address frequently noted Turkish stereotypes prevalent in Germany (Baur & Ossenberg, 2016; Ossenberg, 2019).[6]

[2]See Figures C.8 (p. 258) for an example inbox screenshot, and C.11 (p. 260) for the verified profile badge of a premium user.

[3]For details, see Section C.3 (p. 263) on premium subscription plans for the most relevant online apartment and room rental platforms in selected countries.

[4]All other profile details remain identical. The only variation is the premium status, which is visibly displayed next to the user's full name on the platform.

[5]In 2023, approximately 30% of Germany's population has a migration background, with individuals of Turkish origin constituting the largest ethnic minority (Federal Statistical Office, 2024a, 2024b).

[6]See Chapter 3 for further details, Table B.32 (p. 225) for a list of frequently mentioned stereotypes, and Figures C.12 (p. 260), C.13 (p. 261), C.14 (p. 262), and C.15 (p. 263) for screenshots of the social media profiles.

In a between-subjects, non-matched correspondence test design, each room advertiser is randomly assigned to receive one application from either an ethnic minority or majority male name. The application may or may not include a hyperlink to an Instagram profile, with half of these profiles featuring visual minority stereotypes and the other half not, thereby creating a 2 $\times$ 2 $\times$ 3 factorial design. I conduct the experiment in the 15 largest student cities in Germany over two waves. Both waves were conducted at a similar time of the year to control for demand effects, mitigating demand fluctuations tied to semester schedules as a large fraction of shared apartments consists of students (Kroher et al., 2023). The first wave took place in October and November 2023, shortly after the introduction of the premium subscription service, while the second wave ran from September to November 2024.

A comparative analysis of both waves provides an opportunity to examine how changes in platform design and monetization strategies may affect users' behavior.[7] It also allows for the study of potential first-mover effects that might emerge shortly after the introduction of a new platform feature and then might weaken or disappear over time as the feature becomes more established and market penetration is increased.

The application text features attributes frequently mentioned in applications for rooms in shared apartments, as well as typical student hobbies – aligning with commonly posted Instagram contents (Hu et al., 2014).[8] In the randomly assigned social media conditions, a link to the respective Instagram profile is provided directly in the application which is common for room applications (see Section 2.1). This reduces search costs to virtually zero, as advertisers do not need to search for additional information about the applicant online – in contrast to previous studies (Acquisti & Fong, 2020; Baert et al., 2017; Manant et al., 2019).

All ads posted in the respective cities during the experiment that meet the experimental criteria are contacted. These criteria include: (1) an online time of under three hours, given that landlords typically receive numerous applications within the first few hours – especially in high-demand metropolitan areas; (2) an availability of at least six months, thereby excluding short-term rentals where extensive applicant information may be less relevant; and (3) an age and gender match between the fictitious applicant and the roommate's stated preferences.[9] In addition, professional landlords, such as rental agencies and real estate agents, are excluded, as their information perception and requirements likely differ from those of individual roommates.

A custom program collects all details from each advertisement, such as room and apartment characteristics, advertiser and roommate attributes, and ad statistics, and stores these together with neighborhood- and district-level sociodemographic data in a database.[10] When the fictitious applicant receives a reply from the advertiser or any potential roommate, responses are classified as a *callback* if they involve an invitation to an (online) viewing, call, or meeting; an *other* response if the advertiser requests further information; or a *rejection* if the room is explicitly

[7] One of the very few studies that investigate the effect of changes in platform design and the implications on user's behavior find that a particular change in platform design significantly affected employers' discriminatory behaviors during the hiring process (Manant et al., 2019).

[8] See Section B.4 (p. 212) for a translated version of the application text used in Chapter 3.

[9] Additionally, I exclude ads that are explicitly seeking roommates who are already employed or who want to be contacted via phone, Whatsapp, etc. only.

[10] Additionally, Instagram metrics (e.g., profile visits) are collected manually for each profile before, during, and after the experiment.

denied. To minimize costs to participants (Zschirnt, 2019), any positive reply (callback or other response) is politely declined within 24-48 hours by indicating that the applicant has already found a room on short notice.[11]

4.4 Results

During the first wave, I apply to 2,109 vacant room advertisements, followed by 2,161 applications in the second wave. A total of 58 observations are excluded from the analysis as these involved rental companies (not indicated in the ad but revealed through the response), availability was too short (incorrectly stated ex ante or changed ex post), involved the same landlord (duplicate treatment), or were spam ads that could not be identified as such before the application. Therefore, the final sample consists of 4,212 fictitious applications.

Of these, a total of 1,856 applications (44.1%) received a response. The majority of responses (68.5%) constitute callbacks – invitations to a viewing of the room/the shared apartment and/or a meeting with the potential roommate(s). In 13.5% of responses, the applicant is rejected, and in 18% of responses, the advertiser or roommate(s) request further information about the fictitious applicant.[12]

In the following, I first present results on how information salience affect callback rates (see Section 4.4.1) and discrimination, followed by an in-depth analysis of the degree of salience using a continuous measure, i.e., the position of an application in the advertiser's inbox in Section 4.4.2. While most of the analysis pools both experimental waves, in Section 4.4.3, I analyze differences across the experimental waves to assess potential first-mover effects and a potential change in user's behavior as the premium feature gains traction over time. Section 4.4.4 analyzes heterogeneities in ad quality and salience on the callback probability. Subsequently, Section 4.4.5 shows that social media profiles are actively accessed during the experiment.

4.4.1 Information Salience

Figure 4.1 shows that, an average of 29% of low salience (non-premium) applications receive callbacks, while for high salience (premium) applications this figure rises to 32%, indicating a difference of 3 percentage points. This difference is statistically significant at the five percent level (see Table C.4, p. 248) and indicates that applicants without premium account need to send on average 11% more applications to receive one callback, compared to premium users.

Ethnic minority applicants receive callbacks in approximately 21% of the cases, whereas the respective majority applicants receive callbacks in about 36% when they send their application from non-premium accounts with lower salience. In contrast, in high salience (premium) conditions, these rates are 24% and 39%, respectively. Consequently, both minority and majority applicants exhibit an approximate 3 percentage point improvement in callback rates from higher salience. However, the difference in salience (high versus low) is not statistically significant when examining ethnicities separately (see Panel A of Table C.4, p. 248).

[11] For a detailed examination and discussion of the ethical implications of this research and its experimental design – especially concerning the representation of minority stereotypes – see Section B.8 (p. 236).
[12] See Table C.1 (p. 241) for summary statistics.

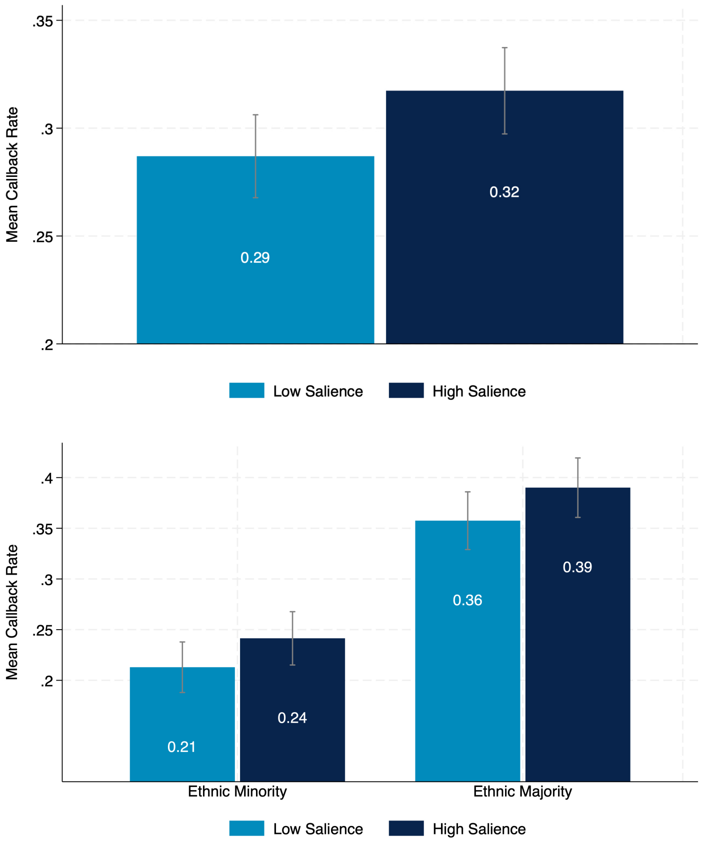


Figure 4.1: Mean Callback Rates for High and Low Information Salience

Regarding the different information conditions, Figure C.1 (p. 247) reveals that an ethnic minority applicant without an additional social media profile (application text only) has a 3 percentage point higher average callback rate (1 pp. for ethnic majority applicants). Again, this difference is not statistically significant (see Panels B and C of Table C.4, p. 248 and Figure C.2, p. 247).

Among applicants who include a social media profile *without* visual minority stereotypes, both applicants gain the most from a higher salience of their application: callback rates increase by six percentage points for minority applicants and seven percentage points for majority applicants, both marginally significant at the 10% level (see Table C.4, p. 248). In percentages, premium status raises callbacks by roughly 26% ($p < 0.1$) for minority applicants and 18% ($p < 0.1$) for majority applicants who do not show minority stereotypes on their social media profile. Similarly, Figure C.2 (p. 247), which plots the average marginal effects of increased salience across the various information treatments (controlling for week, wave, and city fixed-effects), indicates a statistically significant salience effect on callbacks for applicants without minority stereotypes ($p < 0.05$ for minority, $p < 0.01$ for majority applicants).

By contrast, when including a social media profile *with* visual minority stereotypes, a minority applicant, who pays a premium for higher information salience, has a one percentage point *lower* average callback rate compared to a respective non-premium minority applicant. In this information condition, higher salience of stereotypical information even decreases average callback rates for minority applicants. In contrast, the majority applicant has a one-percentage-point higher average callback rate (see Table C.4, p. 248). However, both differences are not statistically significant as confidence intervals are large – albeit smaller for the minority applicant (see Figure C.2, p. 247).

Overall, these descriptive findings indicate that higher salience produces small, mostly insignificant differences in callback rates for most conditions (see Table C.4, p. 248). The largest and marginally significant effects of higher salience appear only for applications that include a social media profile without minority stereotypes. Notably, the salience of an application does not significantly affect discriminatory behavior.

Using different probit models, I further investigate the role of ethnicity, information treatments, and the role of salience controlling for various contextual variables that may affect the probability of receiving a callback. The main specification is as follows:
For a fictitious applicant i, the probability of a callback is given by:

$$\begin{aligned}\Pr(Callback_i = 1) \; = \; \Phi(\alpha \; + \; & \beta_1 \, HighSalience_i \; + \; \beta_2 \, Minority_i \\ + \; & \beta_3 \, (HighSalience_i \times Minority_i) \\ + \; & \beta_4 \, Treatment2_i \; + \; \beta_5 \, Treatment3_i \; + \; \boldsymbol{X}_i \, \boldsymbol{\gamma}),\end{aligned} \tag{4.1}$$

where $Callback_i$ is a dummy variable that equals one if the fictitious applicant receives an invitation to a viewing and/or meeting the potential roommate(s), zero otherwise. $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. $Minority_i$ is a dummy variable that equals one if applicant i has an ethnic minority or ethnic majority name ($Minority_i = 0$). $HighSalience_i$ indicates if the application was sent from a premium account ($HighSalience_i = 1$), zero otherwise. $Minority_i \times HighSalience_i$ captures the interaction between minority and salience.

$Treatment2_i$ (social media profile without stereotypical images) and $Treatment3_i$ (social media profile with stereotypical images) are dummy variables that equal one if the specific information treatment was randomly selected and zero otherwise. $Treatment1_i$, where no social media profile is included in the application, is the reference treatment and therefore omitted. $\boldsymbol{X}_i \, \boldsymbol{\gamma}$ is a vector representing the set of controls, including week, wave, and city fixed effects as well as room/apartment, roommate, and advertiser characteristics, ad statistics, and geographic/demographic variables.

Table 4.1 reports the average marginal effects from different probit models estimating the effects of salience on the callback probability (columns 1 and 2) as well as the effects of salience and ethnicity (Specification 4.1; columns 3 and 4). The models in columns 5 and 6 additionally include interaction terms of salience and the treatment conditions.

Higher salience has a positive and statistically significant effect of 3.5 percentage points ($\beta = 0.035$, $p < 0.01$, see column 2 of Table 4.1) on the probability of receiving a callback. This result confirms the descriptive results when also controlling for week, wave, and city fixed-effects and a large set of contextual variables on apartment, roommate, and advertiser characteristics, as well as geographic and other control variables. Across columns 3 to 6, a minority name has a large and statistically significant negative effect of roughly 15 percentage points ($\beta \approx -0.15$, $p < 0.01$) on the callback probability – even when controlling for a comprehensive set of contextual variables – indicating a large and robust minority penalty. Higher salience still has a positive, statistically significant effect on the callback probability (roughly four percentage points; see columns 3 and 4 of Table 4.1). When interaction terms between high salience and the information treatments are included, the effect decreases and becomes insignificant (see columns 5 and 6 of Table 4.1). Notably, the coefficients of the interaction effects between ethnic minority and high salience are small and not significant across all specifications, aligning with the descriptive results that higher salience (premium status) does not alleviate the ethnic gap in callbacks.

Table 4.1: The Effects of Higher Salience on Callback Probability

Callback	(1)	(2)	(3)	(4)	(5)	(6)
High Salience	0.0338**	0.0347***	0.0361*	0.0393**	0.0229	0.0261
	(0.0138)	(0.0108)	(0.0208)	(0.0153)	(0.0284)	(0.0292)
Ethnic Minority	-	-	-0.145***	-0.146***	-0.144***	-0.145***
			(0.0209)	(0.0183)	(0.0203)	(0.0179)
High Salience × Ethnic Minority	-	-	-0.00410	-0.0104	-0.00545	-0.0120
			(0.0270)	(0.0219)	(0.0263)	(0.0214)
Treatment 1: Without SM (Ref.)	-	-	-	-	-	-
Treatment 2: SM Without Minority-Stereotypes	-	-	0.0469***	0.0460***	0.0190	0.0212
			(0.0128)	(0.0134)	(0.0193)	(0.0136)
Treatment 3: SM With Minority-Stereotypes	-	-	-0.0445***	-0.0495***	-0.0358	-0.0437*
			(0.0136)	(0.0141)	(0.0225)	(0.0225)
High Salience × Treatment 1 (Ref.)	-	-	-	-	-	-
High Salience × Treatment 2	-	-	-	-	0.0548	0.0491
					(0.0356)	(0.0353)
High Salience × Treatment 3	-	-	-	-	-0.0168	-0.0107
					(0.0433)	(0.0393)
Observations	4,212	4,130	4,212	4,130	4,212	4,130
Pseudo R^2	0.032	0.113	0.060	0.146	0.061	0.146
Week & Wave FE	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Room & Shared Apartment Controls	No	Yes	No	Yes	No	Yes
Roommate Controls	No	Yes	No	Yes	No	Yes
Advertiser Controls	No	Yes	No	Yes	No	Yes
Ad Statistic Controls	No	Yes	No	Yes	No	Yes
Geographic Controls	No	Yes	No	Yes	No	Yes
District Level Demographic Controls	No	Yes	No	Yes	No	Yes
P-value Heteroscedasticity (Wald) Test	0.574	0.562	0.256	0.320	0.246	0.344

Note: The table reports the average marginal effects computed from different probit models with callback as the dependent variable. Treatment 1 (no Social Media (SM) Profile included) is the reference treatment and therefore omitted. Columns 1 and 3 report the main effects (treatment conditions) excluding additional control variables while columns 2 and 4 include additional control variables. All specifications include dummies for each week and wave (Week & Wave FE), as well as city dummies (City FE). Additional control variables include room & shared apartment controls (room size, total rent, pictures, online viewing, number of vacant rooms, online time, availability, desired gender and age of a potential roommate, and ad text characteristics), roommate controls (gender of roommates, languages spoken in the apartment, dummies for the type of shared apartment, such as consisting of students, living communally, etc., and other roommate characteristics), advertiser controls (age of the advertiser's account, a profile picture dummy, and advertiser origin), ad statistic controls (views and share of premium applicants), as well as geographic & district level demographic controls (distance to the city center, number of mosques in the area, a dummy for out of town, district population, share of male population, share of Turkish population, and district's population change between 2017 and 2023). Robust standard errors (in parentheses) are clustered at the city level and robust to heteroscedasticity. The Wald test yields $\chi^2 = 0.32$ ($p = 0.574$) for model 1, $\chi^2 = 0.34$ ($p = 0.562$) for model 2, $\chi^2 = 1.29$ ($p = 0.256$) for model 3, $\chi^2 = 0.99$ ($p = 0.320$) for model 4, $\chi^2 = 1.34$ ($p = 0.246$) for model 5, and and $\chi^2 = 0.89$ ($p = 0.344$) for model 6, indicating no significant heteroscedasticity (Heckman, 1998; Neumark, 2012). * p < 0.1, ** p < 0.05, *** p < 0.01.

A social media profile without minority stereotypes (Treatment 2) has a positive effect on the callback probability (compared to the reference treatment), whereas a social media profile with minority stereotypes (Treatment 3) has a negative effect (see columns 3 and 4 of Table 4.1). However, including interaction terms of the information conditions and salience again attenuates the observed effects. In columns 5 and 6, neither profile type generates a significant additional effect, suggesting that increased visibility of the social media conditions with and without visual minority stereotypes does not further amplify or diminish the callback probability – in line with the descriptive results.

When adding triple interaction terms of ethnicity, salience, and the information treatments to the specifications above (see Table C.5, p. 249), the interaction between high salience and social media profiles without minority stereotypes – compared to the reference treatment – becomes large and statistically significant at the 5% level. This result suggests that profiles without minority stereotypes benefit more when an application is highly salient. In contrast, none of the triple interaction terms report significant effects on the probability of receiving a callback. Overall, these findings confirm the descriptive results: profiles without minority stereotypes show the most substantial gains under high salience, but ethnic minority applicants remain at a disadvantage which is not significantly affected by higher salience (premium status) or information treatments.

4.4.2 Inbox Position

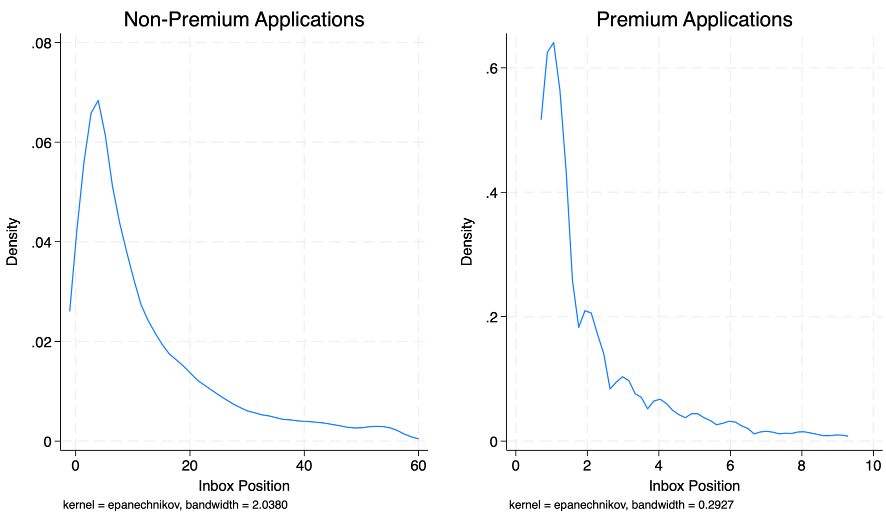


Figure 4.2: Kernel Density Estimates of Inbox Position Distributions

A major advantage of the experimental setting is that it not only permits random assignment of application salience (high versus low levels), but also allows for the direct measurement of an application's precise inbox position at the time it is submitted. On average, the position of the fictitious applications sent from a non-premium account is significantly lower compared to applications sent from premium accounts (16.4 versus 2.9, where a smaller number indicates a higher position in the advertiser's inbox; see Table C.1, p. 241). As Figure 4.2 shows, premium applications are heavily concentrated near the top of the inbox, with a large portion appearing in positions 1 through 3; non-premium applications span a much wider range, often landing between positions 5 and 20 – or even lower. The significant difference in inbox position distributions demonstrates that a premium status substantially increases an application's salience and thus the likelihood of being opened and read by the advertiser.[13]

Table C.7 (p. 252) includes the numerical inbox position to investigate how a lower position in the advertiser's inbox (i.e., a higher value) affect the callback probability. Across all specifications, the linear term for inbox position is negative and highly significant, indicating that the lower (i.e., the higher the number) an application appears, the less likely it is to receive a callback. For instance, an increase of one position reduces the callback probability by approximately 0.27 percentage points (see columns 1 and 2 of Table C.7, p. 252). Although this may appear modest, the cumulative effect can become substantial if the difference spans several positions.

Models 3 and 4 of Table C.7 (p. 252) incorporate a squared term, which is positive and highly significant. This suggests that the negative effect of a lower position flattens out beyond a certain point, indicating the strongest disadvantage occurring between the top and mid-range positions and then decreasing for very low positions.

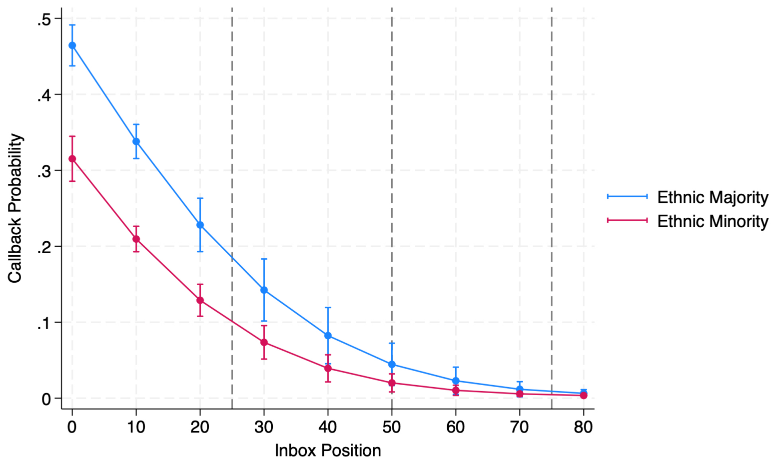


Figure 4.3: Predicted Callback Probability by Inbox Position and Ethnicity

Regarding ethnic differences, Figure 4.3 plots the predicted callback probabilities for both minority and majority applicants, varying the inbox position. Consistent with the regression results, callback rates drop rapidly as the application moves from the top of the inbox down to positions in the 20–30 range, followed by a more gradual decline.[14]

Across the entire range, minority applicants exhibit a substantially lower probability of being called back. While the gap between the two curves diminishes at the lower end of the inbox distribution – where virtually no one receives a callback – it remains substantial across the entire inbox distribution. This underscores the notion that a lower inbox position penalizes all applicants, but disproportionately affects minority applicants – especially on the levels where applications are still likely to be read.

[13]See also Figure C.3 (p. 251) which presents the average number of ad visits and submitted applications – of both premium and non-premium platform users – over time. Even after an entire day, premium applications maintain their prime positions in the inbox compared to non-premium ones.

[14]The platform displays 25 applications per page (indicated by the dashed lines in Figure 4.3).

Moreover, applications from minority applicants appear to be more sensitive to variations in inbox positions. The interaction term between minority and inbox position is negative and statistically significant in all specifications (see Table C.7, p. 252). This implies that being further down in the inbox penalizes minority applicants more than majority applicants.

This result is also evident when examining the information treatments individually. Figures C.4 (p. 253), C.5 (p. 253), and C.6 (p. 254) show the plotted callback probabilities by inbox positions for both minority and majority applicants for each information treatment. In the first condition, without any social media profile (see Figure C.4, p. 253) callback probabilities of minority and majority applicants are significantly different from each other, roughly until the inbox position of 20 (25 applications are displayed per page), until they both flatten out near position 70–80.

When a social media profile free of minority stereotypes is included in the application, the callback probability on the first positions is substantially higher (see Figure C.5, p. 253). However, both probabilities drop quickly and the gap between minority and majority applicants is considerably larger, compared to all other conditions – in line with previous results. This gap is smaller when the applications include a social media profile with minority stereotypes (see Figure C.6, p. 254) where both probability curves flatten out relatively fast (near position 50–60) and the ethnic gap narrows by the bottom of the distribution. In sum, the minority penalty remains evident and statistically significant until around position 20-30 – independent of the information condition.

4.4.3 First-Mover Advantage

Table C.2 (p. 245) compares the first wave (October–November 2023) and the second wave (September–November 2024) across key experimental, response, and ad statistic variables. Response and callback rates decline marginally (from 45% to 43% and 31% to 29%, respectively), suggesting that overall response behavior may have become slightly more competitive. Correspondingly, ad-level data reveal a clear increase in platform activity: listings in the second wave receive significantly more views (136 versus 97), more applications (19 versus 13), and a higher share of premium applicants (rising from around 4.4% to nearly 12%). This suggests that by late 2024, a larger fraction of users had adopted the premium feature, potentially influencing how information salience affect callback rates.

Figure 4.4 indicates that mean callback rates between the waves are not substantially different from another. Across most conditions, higher salience increases callback rates by an average of approximately 3 percentage points. Table C.8 (p. 255) splits the main probit regressions by experimental wave, indicating how higher salience, ethnicity, and social media information affect the callback probability in each wave. The estimates provide mixed evidence regarding a first-mover effect of higher salience for the premium feature. The effects of higher salience are quite consistent across waves producing an average increase in callback probability of about 3.5 percentage points in both waves (see columns 1-4 of Table C.8, p. 255). The negative effect of having a minority name is robust and similar across waves. The consistency suggests that, at least for the initial decision to call back an applicant, there is little evidence of a first-mover effect as the premium feature becomes more established over time.

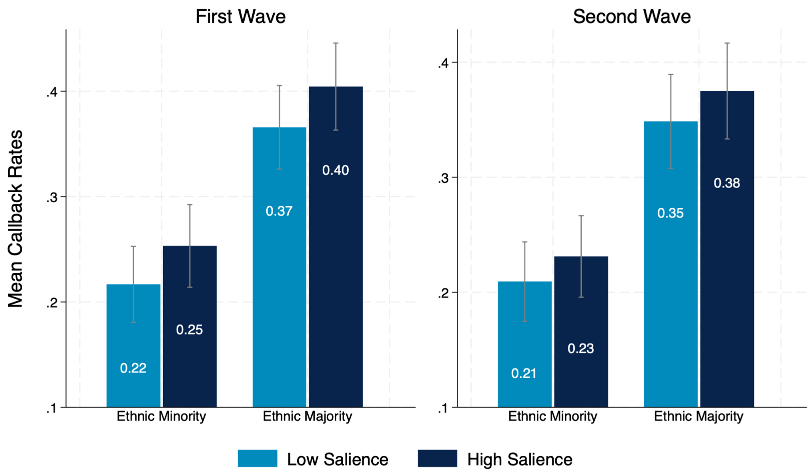


Figure 4.4: Mean Callback Rates per Experimental Wave, Salience, and Ethnicity

The only significant difference between the first and second wave results for applications including a social media profile with minority stereotypes as the direction of the effects change (see columns 5 and 6 of Table C.8, p. 255). Nonetheless, the overall effect (aggregating the individual and interaction effects) of higher salience remains similar in both waves, so that, overall, I do not find evidence for some kind of diminishing advantage from having a premium account over time – irrespective of the information condition. Although market tightness seemed to have increased between the waves, the resulting user behavior does not seem to change with increased adoption of the platform feature.

4.4.4 Ad Quality & Demand Effects

As a proxy for ad quality or demand per ad, I use applications per visit (ratio of applications to visits) of each individual ad. The rationale is that high-quality listings attract more active interest and demand: if a potential roommate visits the ad and finds the offer appealing (e.g., desirable location, good price, appealing photos), she is more likely to submit an application – as opposed to just looking and leaving. Conversely, a low ratio may signal less attractive or over-priced properties.[15]

The models in Table C.9 (p. 256) include applications per visit as a control variable and interaction terms with both ethnic minority and high salience. In both specifications, applications per visit carries a strong negative coefficient ($\beta = -0.978$, $p < 0.01$ and $\beta = -0.685$, $p < 0.01$), suggesting that high-demand or "high-quality" listings are less likely to call back any specific applicant. This is consistent with the notion that in highly competitive ads, each single applicant's probability of a callback decreases. For instance, if a listing accumulates a lot of applications per visit, the advertiser can afford to be selective, decreasing the callback probability for each applicant.

[15] Other factors, such as personal taste or local competition, also play a role in the perceived "quality" of an ad. However, the ratio of applications to visits captures a core component of demand that would otherwise be unobservable.

The minority and salience interaction terms alongside applications per visit do not exhibit statistical significance, indicating no robust evidence that minority applicants or premium applicants with higher salience of their application are disproportionately penalized or favored in high-competition ads. The negative but insignificant triple interaction coefficient ($\beta = -0.249$ in column 1, $\beta = -0.371$ in column 2) implies that while point estimates suggest minority premium applicants may face an additional penalty in highly demanded ads, the data do not rule out the possibility that this penalty is no different from zero. Meanwhile, the main effects of ethnicity and high salience stay broadly consistent with prior results: minority applicants have a roughly 12–13 percentage-point penalty in callback rates, whereas premium status offers a modest but statistically insignificant increase in callback probability once controlling for ad quality.

Overall, high-demand listings have a negative effect on the callback probability of any given applicant, regardless of ethnicity or information salience. Ethnic discrimination is a persistent factor, while salience does not significantly offset the disadvantage – and does not do so more or less strongly as ad competition increases.

4.4.5 Social Media Engagement

Figure C.7 (p. 257) displays daily profile visits over the course of the experiment for the second wave[16]. The Figure indicates that conducting the experiment significantly affect profile visits so that the manipulation is likely to be received by a substantial number of advertisers and/or potential roommates.

Tables C.3 (p. 246) and C.10 (p. 257) present descriptive statistics for social media engagement variables, measured weekly for each Instagram profile condition. These metrics – including visits and impressions – reflect how frequently and intensely users interact with a given profile. In Table C.10 (p. 257), I focus specifically on profile visits and impressions relative to the number of applications sent in a given week. Across all conditions (Panel A), majority profiles achieve roughly 0.25 visits and 1.89 impressions per application, while minority profiles only reach 0.17 visits and 1.32 impressions – both constitute a statistically significant gap at the 1% level. Thus, advertisers or potential roommates generally visit or view majority applicants' social media content more often than the profiles/contents of minority applicants.

Splitting by stereotype condition (Panels B and C) confirms that these gaps persist: majority profiles consistently garner higher engagement, although the size of the disparity and its ratio vary with the presence or absence of visual minority stereotypes. For example, for minority profiles without minority stereotypes (Panel B), the difference in visits is larger (ratio of 1.65) than in the condition with minority stereotypes (Panel C), where the ratio drops to 1.31.

4.5 Limitations

Despite its contributions, this study faces several limitations. Most importantly, I cannot disentangle the effect of salience (higher inbox position) from any perceived stigma around paying for

[16] The respective data for the first wave is not available.

premium status. In particular, a minority applicant – or one perceived to have negative information, characteristics, or others – might appear even more disadvantaged by needing to "purchase" visibility, potentially reinforcing negative inferences about their desirability or resources. While this study treats premium membership as a straightforward enhancement of salience, landlords may interpret it differently, blurring the line between a purely attentional benefit and a more complex "money effect."

Secondly, due to availability, the same social media profiles and accounts were used for both premium and non-premium applications. Thus, it was not possible to observe whether higher salience might alter social media engagement (e.g., profile visits or impressions) for premium versus non-premium users. In other words, the study design does not permit a direct comparison of how being at the top of the inbox might lead to more thorough social media scrutiny. Future experiments with separate premium-only and non-premium-only social media profiles, along with finer tracking of post-application profile views, would clarify whether salience translates into deeper online investigation of certain applicants.

Thirdly, gender was held constant for all fictitious applicants in order to reduce the number of treatment arms. Although this approach isolates the effects of ethnicity and salience, it prevents assessment of how gender interacts with those variables to shape callback outcomes. Future research introducing male and female applicants or other genders would provide a fuller understanding of how multiple social identities affect advertiser's responses.

Lastly, the study is situated in a specific context – the shared housing market in Germany – and focuses on two different ethnic groups within this context. While the findings can, in principle, be extrapolated to other markets and online (matching) platforms characterized by information asymmetry, future research might explore the effects of information salience in different markets, platforms, and settings – and whether and to what extent information salience can affect (statistical) discrimination.

4.6 Discussion & Conclusion

The effects of information salience on human behavior has attracted growing attention in economics and social sciences in the last decade (e.g., Bordalo et al., 2022; Frydman & Wang, 2020; Ghose et al., 2014). However, a large fraction of the existing literature focuses on the effects of salience on consumer behavior (Bordalo et al., 2012, 2013; Chetty et al., 2009; Michels et al., 2024).

The extent to which information salience affects labor or housing market outcomes, i.e., callback rates, and how it affects ethnic discrimination is largely unknown (Bordalo et al., 2022; Manant et al., 2019). Given the increasing importance of online matching platforms for a myriad of purposes, including job or apartment search, dating, friendships, and more, this paper addresses a significant research gap in the experimental literature on the role of information salience of applications, requests, or inquiries on (online) matching and market outcomes. Furthermore, this paper investigates the interplay of salience and discrimination, and extends the existing literature on how changes in platform design affect market outcomes of disadvantaged groups (Cheng et al., 2024; Doleac & Stein, 2013; Lambrecht & Tucker, 2019).

According to salience theory (e.g., Bordalo et al., 2012, 2022), attributes that stand out in a decision-maker's environment disproportionately affect outcomes. When salient cues highlight a person's minority identity – whether through a name, photograph, or social media profile – stereotypes and other beliefs may be triggered more readily, intensifying discriminatory behavior. Conversely, where salient information counters stereotypes or presents the applicant in a favorable light, it may offset existing biases.

The findings of this study confirm the general intuition that higher information salience – operationalized through a premium account yielding substantially higher inbox positions – can significantly improve overall callback rates by approximately 3 percentage points. This result is generally in line with the theoretic prediction that reducing reading costs increase the likelihood of an application being considered. Having an application near the top in a potential roommate's, landlord's, or advertiser's inbox appears to capture increased attention and translates into a higher callback probability.

Importantly, the increase in callback probability due to higher salience is not affected by ethnicity. The average ethnic gap remains constant between non-premium and premium applications. This finding is inconsistent with the theoretic prediction that minority applicants gain less from higher salience compared to majority applicants. Furthermore, it is somewhat inconsistent with Manant et al. (2019), who document a 42% callback gap between applicants whose ethnic origin (indicating a minority vs. majority origin) is only available on social media – yet observe that this gap fades away once the salience of the foreign-origin signal is reduced due to an exogenous change of the social media platform's layout. However, the experimental design of Manant et al. (2019) did not allow for a random variation of salience.

In addition, analyzing the precise numerical inbox position of each application reveals that a one-position-decrease in rank reduces the probability of receiving a callback by approximately 0.3 percentage points. Although this effect may seem modest, it can accumulate considerably across multiple positions. While lower inbox positions penalize all applicants, they disproportionately affect minority applicants – particularly in the range where applications are still likely to be read.

Moreover, the significant ethnic gap between minority and majority applicants persists, irrespectively of the additional social media information that advertisers or roommates are presented with. The results indicate that stereotype-challenging cues (social media profiles without minority stereotypes), when presented with higher salience, can help minority applicants achieve slightly higher callback rates. However, the same is true for majority applicants whose profile is free of minority stereotypes. Thus, these results are also inconsistent with the theoretic prediction that minority applicants lacking stereotypical cues on their social media profiles gain a greater boost from high-salience placements compared to majority applicants.

By contrast, callback rates of minority applicants providing stereotypical information *and* having a premium account are even marginally *lower* than non-premium minority applicants with the same information. Although this difference is not statistically significant, it shows that higher salience may actually be penalized in the case of information that is potentially perceived negatively or that confirms or reinforces stereotypes. This patterns support the theoretic prediction that increased salience can amplify the impact of identity-related cues so that unfavorable signals exacerbate biases.

At the same time, the ethnic gap remains robust, particularly in high-demand listings where competition is fierce. Consistent with prior research on market congestion, greater competition lowers the probability of receiving a callback for all applicants, irrespective of ethnicity or information salience, yet the minority penalty remains again unaffected. I do not find a first-mover effect for premium adopters, when comparing the results of both experimental waves.

It is striking that the social media information affects user behavior: engagement data reveal that landlords or potential roommates indeed visit the linked profiles at significant rates. This confirms that information about an applicant's background or personal interests (whether stereotype-consistent or not) bear the opportunity to shape decision-making (see also Chapters 2 and 3 of this thesis).

By design, this study treats premium status purely as a conduit for higher salience. However, part of the effect attributed to salience could in fact stem from a "money effect" that prompts inferences about socioeconomic status, overshadowing or compounding any straightforward attention-based benefit. In particular, potential roommates may infer that a minority applicant is "desperate" or otherwise disadvantaged if they must pay for better positioning, possibly entrenching negative stereotypes. Future research could investigate how these perceptions interact by separating a purely attentional boost from stigma related to paying for salience.

Overall, the experiment shows that increasing salience by moving applications towards the top of the inbox slightly improves callback rates. However, it does not narrow the ethnic gap. If anything, the resulting attention appears to magnify the importance of whatever additional information about minority status is provided, for better or worse, and minority applicants who post stereotypical minority cues even see a negative benefit from increased visibility.

These insights hold particular relevance for platforms that introduce design changes or monetization strategies that might either mitigate or intensify discriminatory behavior and preferences. In the end, raising salience is not a panacea: it can increase overall responsiveness, but it does not eradicate biases. From a user's perspective, paying a premium may not be worth the comparatively small additional gain of higher salience of an application, especially not if a minority user wants to increase her callback chances compared to majority applicants.

Consequently, policy measures aimed at reducing discrimination may need to go beyond advising users to carefully consider what to post on social media. For example, anonymized or standardized screening methods in which personal or origin-related cues are masked can prevent recruiters or landlords from relying on such cues altogether. In addition, stricter anti-discrimination enforcement and greater transparency requirements could deter biased decision-making. From a platform-design perspective, implementing layout features that de-emphasize or reorder personal attributes – or actively remind decision-makers to consider relevant qualifications or characteristics – could further discourage discriminatory screening behaviors. By proactively shaping how information is presented, platforms can help ensure that (paid or subscription-based) features do not inadvertently exacerbate ethnic disparities, and instead foster fairer outcomes for all users.

Chapter 5

Causal Effects of Personality Signals on Online and Offline Outcomes[1]

5.1 Introduction

The use of social media platforms and their role for social network formation, interactions, and online self-representation have reached unprecedented levels (Back et al., 2010; Datareportal, 2024; Evsyukova et al., 2025) with social media fundamentally transforming the ways in which individuals interact and evaluate information (Allcott et al., 2020; Aridor et al., 2024). Social media platforms provide rich insights into users' personality traits (Back et al., 2010; Gosling et al., 2007) which play an important role in predicting individual outcomes in different contexts (Barrick & Mount, 1991; R. Fang et al., 2015; Judge et al., 1999; Rothmann & Coetzer, 2003). While several studies have examined how self-, friends-, and system-generated information on social media profiles affect impression formation (Utz, 2010; Walther et al., 2009; Walther et al., 2008), it has not been analyzed as yet how personality cues on social media profiles affect real-world outcomes.

We address this gap by investigating the effects of social media personality cues on real-world outcomes. Specifically, we conducted two randomized large-scale field experiments in two different settings, and causally identified the effects of social media personality cues on real-world decisions. In our interdisciplinary endeavor, we effectively combined two different research traditions: economists' experience in conducting randomized large-scale field experiments and communication scholars' insights on the role of social media personality cues.

To conduct our experiments, we created four fictitious social media profiles on the photo sharing platform Instagram and varied visual cues to signal the personality traits of the fictitious profile owners. In Experiment 1, we sent friend requests from our fictitious profiles to other users, as suggested by Instagram's algorithm, and measured acceptance and re-follow rates. In

[1] Chapter 5 is based on the working paper "How Social Media Personality Cues Affect Selection Decisions: Evidence From Two Large-Scale Field Experiments" by Raphael Moritz (University of Tübingen), Kerstin Pull (University of Tübingen), and Sonja Utz (University of Tübingen and Leibniz-Institut für Wissensmedien). The implementation of the field experiment was approved by the Ethics Committee of the Faculty of Economics and Social Sciences of the University of Tübingen (A2.5.4-133_aa) on September 25, 2020. The research project was pre-registered in the AEA RCT Registry (#12473 and #13359): https://doi.org/10.1257/rct.12473-1.0 and https://doi.org/10.1257/rct.13359-1.0.

Experiment 2, we used the same fictitious profiles, linked these to applications for vacant rooms ads on a platform for shared apartments and measured callback rates. In addition, we also applied for vacant rooms without linking the application to an Instagram profile – thus being able to additionally assess the effects of social media cues compared to a situation where *no* social media information is provided.

Our research offers several contributions. Firstly, we contribute to the literature on impression formation on social media. While much of the work on impression formation was conducted with (fictitious) Facebook posts (Tong et al., 2008; Utz, 2010; Walther et al., 2009), we focused on Instagram, a platform that is much more dominated by pictures. Secondly, and more importantly, instead of letting participants evaluate the screenshot of a profile in a hypothetical setting, we conducted field experiments where participants were able to assess complete profiles and then took *real* decisions with real-world consequences (i.e., whether to accept someone's friend request or invite someone for a room tour), thus studying settings with high external validity. To make sure that our fictitious Instagram profiles were realistic in terms of friends and posting frequency, we had prepared these over years. While prior work looked, at best, at behavioral intentions (Domahidi et al., 2022; Utz, 2010), we examined responses in terms of real-world decisions. Thirdly, we studied the real-world effects of social media cues in two different settings that varied in the severity of the consequences of the decisions taken: Experiment 1 represents a low-stake setting in terms of the consequences being faced by both, the profile owner whose friend request has or has not been accepted and the user who decides on accepting the friend request or re-following the account. In comparison, the setting of Experiment 2 is of considerable higher stake – for both, the person who searches for a room and the incumbent roommates who invite that person to their apartment and who might even share their apartment with that person in the future. By studying whether the effects of personality cues vary across the two settings, we contribute to the growing strand of research that examines how online signaling relates to offline outcomes (Moritz et al., 2023; Stoughton et al., 2013; van Zoonen & van der Meer, 2015). Lastly, by manipulating personality traits via social media cues, we explicitly link three of the Big Five personality traits (McCrae & Costa, 1987) to real-world outcomes.

5.2 Impression Formation on Social Media

Research on impression formation on social media falls into three categories: (1) studies examining the accuracy of impressions by correlating self-ratings with the ratings of external observers; (2) studies measuring social media users' personalities and analyzing – manually or computationally – their profiles, and (3) experimental work examining the influence of different cue types on judgments.

The first line of research addressed the question of how accurate the self-generated profile information is and whether it can be used to form a reliable impression of a person. Most researchers focused on the Big Five which are considered as the core personality traits: openness to experience, conscientiousness, extraversion, agreeableness and neuroticism (or emotional instability) (McCrae & John, 1992). The Big Five substantially and significantly affect almost all domains of life outcomes, ranging from social relationships to academic and job performance (Asendorpf & Wilpers, 1998; Barrick & Mount, 1991; Barrick et al., 2001; Poropat, 2009). Although people tend to present themselves positively on social media platforms, several studies found that profiles are largely accurate in terms of displaying a profile owner's personality, and show only a small degree of idealization (Back et al., 2010; Toma & Carlson, 2015; Toma & Hancock, 2013). Positive correlations between self- and observer ratings were found in several studies (Back et al., 2010; Gosling et al., 2007; Stopfer et al., 2014).

A second line of research examined the cues that are used to form an impression. Work in this domain focused on profile information and posts (text and photos). Some scholars manually coded the profile information and found distinctive links between cues and personality traits (Hall et al., 2014; Stopfer et al., 2014). Hall et al. (2014), for example, found that the number of friends, the number of friends on pictures, or status updates with positive affect were related both to self-reported extraversion and the extraversion judgments made by others. Other researchers analyzed profiles with machine learning algorithms. This work not only focused on the content of texts or photos but also considered characteristics like language use and attributes of the photos such as colors, color diversity, saturation, and settings. Yarkoni (2010), for example, presented a list of words most frequently used by people scoring high on the respective Big Five trait. Cucurull et al. (2018) and Segalin et al. (2017) report that images exhibiting high levels of agreeableness are conveyed through bright, warm, and colorful scenes – often showing nature, water, or flowers – while low levels of agreeableness manifest as black-and-white or heavily desaturated images, with cluttered or text-laden backdrops and fewer friendly elements (Segalin et al., 2017). Pictures that signal high emotional stability frequently incorporate natural scenery with warm colors, while low emotional stability is associated with grayish photos with shadowed faces (Cucurull et al., 2018). Images signaling high levels of conscientiousness often feature healthy foods such as vegetables or depict exercise-related content (Cucurull et al., 2018), while photographs signaling lower levels of conscientiousness often rely on faded or pastel colors, gray or monochromatic backgrounds, and sparse color variation (Segalin et al., 2017).

The third line of research investigated the underlying processes. Several studies build on warranting theory and compare the impact of self- and other-generated information (Tong et al., 2008; Utz, 2010; Walther et al., 2009; Walther et al., 2008). The central assumption of warranting theory is that information which is more difficult to manipulate has a higher warranting value and, consequently, a larger impact on impression formation. Other-generated information can be information created by friends (e.g., comments on one's posts or pictures) or the system (e.g., the number of friends). These studies used experiments with mock profiles and examined, for example, the effect of the number of Facebook friends on popularity judgments (Tong et al., 2008) or contrasted the effects of self-generated posts and comments by friends on the attractiveness or extraversion of a target (Utz, 2010; Walther et al., 2009). Although other-generated information has an impact, warranting theory was only partially supported with self-generated information having large effects on impressions.

Taken together, we know that self-presentation on social media profiles is fairly accurate, that people are able to form accurate impressions from social media profiles, and that they do so also by relying on self-generated information. Additionally, several studies provide information on the textual and visual cues that are related to personality traits. However, it remains unclear whether these impressions affect real-world decisions as most research was conducted using, at best, behavioral intentions as outcomes (Domahidi et al., 2022; Tong et al., 2008; Utz, 2010; Walther et al., 2009; Walther et al., 2008).

5.3 The Present Study

The aim of the present research is to examine the influence of social media personality cues on real-world decisions in externally valid field settings. To do so, we brought together two research lines. The work on impression formation has shown that people can form accurate impressions based on social media profiles. However, this work only looked at judgments and did not address behavioral outcomes. Personality psychologists have explored the relationships between the Big Five and a variety of outcomes. What is missing is work that demonstrates in large-scale field experiments that impressions built on the basis of social media profiles affect real-world outcomes.

Our first overarching research question is:

RQ1: What role do personality traits, as displayed through social media profiles, play in the decision to accept a friend request (Experiment 1) or to call back an applicant (Experiment 2)?

We selected these two outcomes and corresponding settings because they vary in the severity of their consequences: The decision to accept a friend request on Instagram is, compared to inviting that someone to your apartment as a potential roommate, rather low stake. Further, accepting a friend request only takes seconds – as does ending a virtual friendship. In contrast, when sharing an apartment, issues such as cleaning, washing dishes, or noise levels can create major conflicts. In addition, parting with a roommate may be difficult and costly.

In our study, we focus on three of the Big Five traits: conscientiousness, agreeableness, and emotional stability (see also sections D.5.1, p. 296, D.5.2, p. 301, and D.5.3, p. 305). Conscientiousness refers to the extent to which an individual is efficient, organized, planning, reliable, responsive, and thorough (McCrae & John, 1992). Agreeableness refers to the degree to which a person is grateful, forgiving, generous, kind, sympathetic, and trusting (McCrae & John, 1992). Emotional stability, conceptualized as the inverse of neuroticism, refers to an individual's ability to remain calm, composed, and resilient in the face of stress and adversity (McCrae & John, 1992). High emotional stability is associated with effective emotional regulation, reduced susceptibility to negative emotions, and the capacity to maintain positive interpersonal relationships even under pressure (John et al., 2008).

We focus on these traits, because our first online survey that we conducted in preparation of Experiment 2 (see Section D.5.1, p. 296) suggested that individuals prefer roommates characterized by high agreeableness, emotional stability, and conscientiousness. With regard to social network formation (our setting in Experiment 1), several studies hint at agreeableness, emotional stability and conscientiousness being linked to the formation, maintenance, or quality of friendships. For instance, Harris and Vazire (2016), Jensen-Campbell et al. (2002), and Selfhout et al. (2010) show that agreeable people are more often selected as friends. While the review by Harris and Vazire (2016) reports no effects of emotional stability on friendship formation, it does so for relationship maintenance. Low emotional stability is related to more conflicts and lower relationship satisfaction for both, the person scoring low on emotional stability and the friend (Harris & Vazire, 2016). For conscientiousness, Harris and Vazire (2016) report little or no effect on becoming friends, but on relationship quality. Jensen-Campbell and Malcolm (2007) found that adolescents who score high on conscientiousness report higher peer acceptance.

Concerning the effects of agreeableness, emotional stability and conscientiousness on our outcome variables in the two experimental settings, we expected to observe the same patterns. Because we were not able to independently manipulate agreeableness and emotional stability (see Section 5.4 below), we combined these two traits. To simplify language, we speak of high vs. low agreeableness/emotional stability instead of distinguishing between high agreeableness/low neuroticism and low agreeableness/high neuroticism. The hypotheses for the roommate setting in Experiment 2 (H1b and H2b) were preregistered. Below, we expand these to also include the effects for the friend request setting in Experiment 1, which we expect to be completely analogous to the roommate setting in Experiment 2 (H1a and H2a).

H1: People with social media profiles bearing traits associated with high agreeableness/emotional stability are more likely to (a) be accepted as Instagram friend and (b) receive callbacks than those with low agreeableness/emotional stability.

H2: People with social media profiles bearing traits associated with high conscientiousness are more likely to (a) be accepted as Instagram friend and (b) receive callbacks than those with low conscientiousness.

In Experiment 2, we were able to include a treatment in which no social media information was provided and thus address a second research question:

RQ2: What is the impact of personality traits as displayed through social media profiles on the decision to call back an applicant – compared to *no* social media information being provided?

By comparing the effects of signaling a certain level of agreeableness/emotional stability or conscientiousness to not providing any social media information, we intended to learn more about the drivers of the effects studied in RQ1. Further, we sought to assess whether providing social media information does affect real-world decisions – irrespective of its content. Assuming that not providing any social media information would serve as a neutral baseline and that additional information from linking a profile to the application for a room in a shared apartment would be used to adjust this default impression, we predicted the effects described below. The corresponding hypotheses were preregistered.[2]

H3: Applicants with social media profiles signaling high agreeableness/emotional stability are more likely to receive callbacks compared to applicants without additional social media information (baseline treatment).

H4: Applicants with social media profiles signaling high conscientiousness are more likely to receive callbacks compared to applicants without additional social media information (baseline treatment).

[2] In the pre-registration, H3 and H4 were originally combined in one hypothesis and H5 and H6 in another. To enhance readability, we have separated them into four distinct hypotheses.

H5: Applicants with social media profiles signaling low agreeableness/emotional stability are less likely to receive callbacks compared to applicants without additional social media information (baseline treatment).

H6: Applicants with social media profiles signaling low conscientiousness are less likely to receive callbacks compared to applicants without additional social media information (baseline treatment).

5.4 Creating Stimuli For Our Experiments

We created four fictitious social media profiles on the social media platform Instagram using pictures that signal high and low levels of agreeableness/emotional stability and conscientiousness. All images featured the same 24-year-old female student, who was recruited via a campus-wide email and subsequently selected by students from a candidate pool based on the female student's social media appearance. We chose "Julia" as the profile's first name – one of the most common female names among the alleged profile owner's birth cohort – and paired it with the last name "Becker" – one of the most prevalent German surnames (Gesellschaft für deutsche Sprache, 2019; Kohlheim & Kohlheim, 2005).

We incorporated data on biographical information typically found on Instagram profiles of university students – such as age, occupation, and hometown – and incorporated these into Julia's profile. Drawing on data indicating that the median age of students in Germany is approximately 23.5 years (Federal Statistical Office, 2024a), we assigned a fictitious age of 24. We then indicated "student" (accompanied by a small graduation cap emoji) as her occupation, and designated Tübingen, known for its large student population, as her alleged hometown.[3]

We registered the profiles in August and September 2019. To simulate a natural progression of content, we initially uploaded some neutral images over several months, depicting nature and travel scenes without featuring any individuals.

We then first compiled sets of images illustrating high and low expressions of agreeableness and conscientiousness, in line with the text- and picture-based approaches described above (see Section 5.2, p. 82). To verify whether the selected images effectively manipulate individuals' perceptions, we conducted a first randomized online pilot experiment with $n = 467$ students (see Section D.5.2, p. 301). In a between-subjects design, participants were randomly assigned to one of four treatments (high agreeableness, low agreeableness, high conscientiousness, low conscientiousness) and then asked to rate the perceived personality of the individual using the Big Five Inventory-2-S (Rammstedt et al., 2020; Soto & John, 2017) on a 5-point Likert scale. While the results of this first pilot experiment confirmed that the conscientiousness manipulation was successful, they did not demonstrate comparable effects for the agreeableness manipulation (see Section D.5.2, p. 301).

Consequently, we refined the images for high vs. low agreeableness and conducted a second randomized online pilot experiment with $n = 652$ participants (see Section D.5.3, p. 305). We

[3] In Experiment 2, the application text for the room in the shared apartment further specified that Julia studies business and economics (which are the most common fields of study (Federal Statistical Office, 2024b), and that she is a Master level student.

carefully designed the captions that describe a photo's contents to be in line with YARKONI (2010) – thus ensuring that the captions accurately reflected the respective treatment[4]. In this second pilot experiment, the conscientiousness and the agreeableness manipulations were successful.[5] However, because emotional stability showed similar patterns as agreeableness (see Tables D.27, p. 309, and D.28, p. 309), we consolidated the high vs. low agreeableness treatments as high vs. low "agreeableness/emotional stability".

The final high and low conscientiousness profiles each featured 35 posts, the high agreeableness/emotional stability profile featured 36 posts and the low agreeableness/emotional stability profile featured 30 posts. All profiles featured one story item with images from the alleged hometown of the profile owner.

In an additional online survey we asked students to assess the authenticity of the created profiles. 89.4% of the respondents considered the profiles to be at least somewhat realistic, with 60.6% rating them as extremely or very realistic. Only a small fraction (3%) reported finding the profiles unrealistic (see Table D.32, p. 314, in Section D.5.5, p. 312).

Besides having survey participants rate a profile's authenticity, we also asked them to subscribe to the respective profile and like a random selection of pictures[6] (see Section D.5.5, p. 312). Firstly, the presence of a social network reduces the likelihood of the corresponding profiles being perceived as fictitious or inactive by the participants in the later experiments. Secondly, by randomly assigning students to profiles, we established an evenly distributed social network that serves as the basis for the platform's friend-suggestion algorithm which we exploited in the friendship setting in Experiment 1. Thirdly, maintaining a network of genuine users reduces the likelihood that our profiles would be flagged as fictitious by the social media platform.[7] In addition, a subset of survey participants was also asked to subscribe to additional private accounts that we had created. These accounts were non-public, depicted students, and served as placeholders for friends, allowing us to tag them in photos on the primary experimental profiles where peers are visible.

During Experiment 2, the profiles had an average of 145 (125 – 165) followers and 304 (296 – 311) subscriptions. The majority of followers was between 25 and 34 years old (56.6 – 70.5%) and indicated being male (52.8 – 59.1%).

All profiles were set as public so that potential friends and roommates could view the entire profile at basically no cost.[8] Furthermore, we disabled the option to suggest similar profiles on both, experimental and additional profiles, to avoid contamination (see Figure D.19, p. 296).

[4]Examples include: "Healthy food to go with fruits, vegetables and delicious wraps #healthy #snacktime #enjoy #foodie #delicious #healthyfood #tasty" for a food image on the high conscientiousness profile or "frozen pizza4eva!! YOLO!!!1!! #noregrets #protest #sorrynotsorry" for a food image on the low conscientiousness profile.

[5]See Figures D.12 (p. 291), D.13 (p. 292), D.14 (p. 293), and D.15 (p. 294) for screenshots of the profiles/treatments.

[6]To simulate a profile's evolution over time, more recently posted images were more often displayed when subjects were asked to like photos.

[7]Throughout the experiments, we monitored whether any of the profiles was "flagged for review" – a measure Instagram employs to detect spam or fraudulent accounts. We did not find any instances of such flags being applied.

[8]As a result, also non-registered users may view the entire profile page, including all posts. To access additional information such as the lists of friends and subscribers, however, a login is required.

5.5 Experiment 1

5.5.1 Method

Experiment 1 had a 2 (personality trait: conscientiousness vs. agreeableness/emotional stability) x 2 (level: low vs. high) design. Design, procedure, measures, and planned sample size were preregistered in the AEA RCT Registry (#13359). We conducted the experiment in June 2024.

We started by collecting all suggested friends for each of the four Instagram profiles with all profiles featuring the same individual, including the same name, age, student status, and alleged hometown (see Section 5.4, p. 86).[9] When a particular user was suggested to more than one of the profiles, we randomly chose one of the four treatment profiles to send a friend request to the respective user.[10] As a result, each suggested user received only a request from one of the profiles to reduce the likelihood of detection and the costs associated with involuntary participation in the experiment.[11]

We excluded verified profiles in the subject pool because they primarily belong to persons of public interest, celebrities, or influencers, who typically do not respond to requests and do not follow back.[12] Business and non-profit profiles were not excluded, but only accounted for 2% of the suggested connections.

We included both private and public profiles. The visibility settings of a profile do have important implications for the experimental outcomes: private accounts require approval (or denial) for any friend request, whereas public profiles automatically add requesting users as new subscribers to their follower lists.[13] Both, users with private and public profiles may re-follow an account.

We wrote a computer program that collected friend suggestions, managed subject pools and sent friend requests. Upon receiving a friend request, subjects see a notification displaying the requester's profile picture, username, and full name.[14] Furthermore, the program collects and stores all publicly accessible profile data – including the suggested user's displayed name, biographical information, number of posts, subscribers and subscriptions, mutual connections (if any), and others – in a database. Note that we did not collect any data that only became available after a suggested user had accepted a friend request.

[9] See Figure D.16 (p. 295) for a screenshot illustrating the platform's suggestion feature.

[10] Due to technical constraints, we could not fully randomize the assignment of potential friends to profiles: the pool of suggested friends remained relatively stable over time if no friend requests were actually sent. Thus, after each round, the number of new suggestions would have been decreased so that the experiment would not have been feasible. Nonetheless, as the suggestions were based on each profile's existing social network – predominantly comprising students randomly assigned to those profiles (see sections 5.4, p. 86, and D.5.5, p. 312) – we argue that the sample selection is quasi-random. Additional robustness checks (see Section 5.5.3, p. 90) indicate that interaction effects of the treatment variables and heterogeneities in friend suggestions do not exhibit statistically or economically significant results in the majority of cases.

[11] See Section D.6 (p. 314) for a short discussion on the ethical implications of our research.

[12] We further excluded all usernames who participated in the pilot studies, research assistants involved in the experimental design, the female student who provided us with her photos and her subscribers, the subscribers of the experimental profiles and subscriptions of these (including peripheral profiles) to prevent that people who were aware of the experiment from being treated. In so doing, we avoided the risk of contamination.

[13] However, users with public profiles may subsequently remove subscribers if desired.

[14] Figures D.17 (p. 295) and D.18 (p. 296) present example screenshots (desktop and mobile app views, respectively). The examples showcase all treatment accounts. Note that in the experiment, any user receives a single request only.

In total, we sent friend requests to 1,002 users suggested by Instagram's algorithm.[15] Approximately 65% of the suggested profiles were private, with an average of 36 posts, 691 followers, and 528 subscriptions (see Table D.1, p. 266). On average, each suggested friend shared four mutual connections with our experimental profiles. 48% identified as female. The mean age of suggested friends was 26 years. Among profiles featuring a biography, 6% report living in Tübingen and 12% indicate being a student.

Our primary dependent measure was the *acceptance rate* (for private profiles) or the *re-follow rate* (for private and public profiles) tracked by the computer program. After sending a friend request, the program periodically revisited the respective profile. If a user reacted – whether by accepting and/or re-following – the program recorded the reaction as well as any newly formed mutual connection. This allowed us to address potential endogeneity – such as a user's decision to accept a friend request because a mutual connection had done so meanwhile – thus enhancing the robustness of our findings (see Section 5.5.3, p. 90).

Furthermore, we collected data on profile visits, impressions, and other engagement metrics for each profile over the course of the experiment.

5.5.2 Results

Figure 5.1 reports the mean acceptance rates for each of the four experimental conditions. The difference between high and low agreeableness/emotional stability (46% vs. 29%) is statistically significant ($z = -3.080$, $p = 0.002$), as the results of a two-sample Wilcoxon rank-sum (Mann–Whitney) test indicates (see Table D.2, p. 267, for further details). This 17-percentage-point gap implies that profiles signaling low agreeableness/emotional stability would need to send approximately 57% more friend requests than their high level counterparts to receive the same number of acceptances. Hence, H1a is supported. In contrast, we find no support for H2a: Although high levels of conscientiousness are associated with higher acceptance rates than low levels of conscientiousness (42% vs. 39%), this difference is not statistically significant ($z = -0.568$, $p = 0.648$).

Concerning re-follow rates (Figure 5.2), we find further support for H1a and again no support for H2a: Profiles indicating high agreeableness/emotional stability elicited a 20% re-follow rate, while low agreeableness/emotional stability profiles generate only an 8% re-follow rate, resulting in a significant gap of 12 percentage points ($z = -3.900$, $p = 0.000$; see Table D.2, p. 267). That is, profiles indicating low agreeableness/emotional stability needed to follow 153% more users to receive the same number of re-follows compared to high agreeableness/emotional stability profiles. By contrast, profiles signaling high levels of conscientiousness received re-follows in 18% of cases, compared to 23% for low-conscientiousness profiles – a difference that is descriptively in the opposite of the predicted direction, but not statistically significant ($z = 1.370$, $p = 0.207$).

Additionally, we conducted regression analyses including control variables available from the profiles. Table 5.1 reports average marginal effects from different (heteroskedastic) probit models predicting acceptance and re-follow probabilities. For acceptance (columns 1 and 3), high

[15] See Figures D.17 (p. 295) and D.18 (p. 296) for screenshots of the treatment notifications.

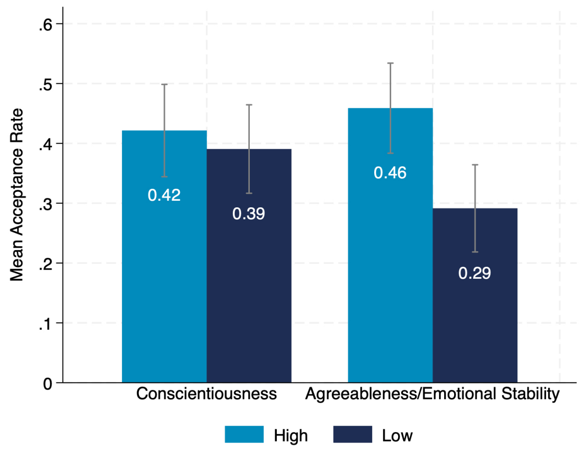


Figure 5.1: Mean Acceptance Rates

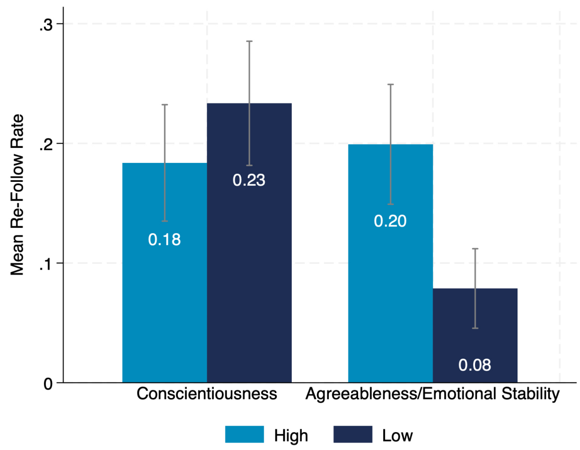


Figure 5.2: Mean Re-Follow Rates

levels of agreeableness/emotional stability show large and statistically significant positive effects on the likelihood of the friend request being accepted ($\beta = 0.116$, $p = 0.008$; see column 3) – compared to low agreeableness/emotional stability. High levels of conscientiousness exhibit a small positive effect on the probability of the friend request being accepted compared to low levels of conscientiousness, which is, however, not statistically significant ($\beta = 0.016$, $p = 0.737$; see column 1). The probability of being re-followed (columns 2 and 4) follows suit with a positive and statistically significant effect for agreeableness/emotional stability being high vs. low ($\beta = 0.072$, $p = 0.020$; see column 4) and a non-significant small negative effect for conscientiousness being high vs. low ($\beta = -0.030$, $p = 0.396$; see column 2). Our regression results thus confirm the above findings – controlling for a comprehensive set of characteristics of the suggested friends' profiles.[16]

5.5.3 Robustness Checks

Endogeneity by Mutual Connections Potential endogeneity may arise if a user's decision to accept a friend request is affected by another mutual connection that accepted the request in the meantime. To address this concern, we assessed the impact of changes in mutual connections (between sending the request and the reaction) on the probability of accepting friend requests and re-following profiles, as presented in Table D.6 (p. 272).

Changes in mutual connections do not generally seem to be a significant driver of acceptance or re-follow probabilities at conventional levels of significance. While some of the interaction terms in the models with the re-follow probability as the outcome for high conscientiousness (see column 2 of Table D.6, p. 272) and low agreeableness/emotional stability (see column 4) achieve

[16]One of these user characteristics, being female, has a large negative and statistically significant effect on both, the acceptance and re-follow probability – hinting at females being more selective when it comes to accepting or re-following accounts of people they do not know. Estimating a similar model as presented in Table 5.1 including additional biographic variables, such as location(s), student status, or motto/saying in the bio (see Table D.3, p. 268) – indicates that most of these additional contextual variables do not significantly predict acceptance or re-follow probabilities. Only one variable, i.e., having a motto or quote in one's bio, does have an effect and is associated with a lower likelihood of accepting a friend request (see columns 1 and 3 of Table D.3, p. 268).

Table 5.1: Determinants of Reaction Rates (Average Marginal Effects)

	(1)	(2)	(3)	(4)
	Conscientiousness		Agreeableness/ Emotional Stability	
	Accept	Re-Follow	Accept	Re-Follow
Treatment: Low (Ref.)	-	-	-	-
Treatment: High	0.016	-0.030	0.116**	0.072*
	(0.047)	(0.035)	(0.044)	(0.031)
Posts	-0.001	-0.000	-0.001	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Followers/subscribers	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Following/subscriptions	0.000	0.000**	0.000	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Mutual connections	0.004	0.005†	0.008	0.006*
	(0.005)	(0.003)	(0.005)	(0.003)
Account age (months)	0.001	-0.000	0.000	-0.000
	(0.001)	(0.000)	(0.001)	(0.000)
Bio	0.010	-0.006	0.010	-0.014
	(0.044)	(0.027)	(0.043)	(0.028)
Female	-0.231***	-0.183***	-0.233***	-0.180***
	(0.034)	(0.025)	(0.034)	(0.024)
Private		0.039		0.040
		(0.027)		(0.027)
Obs.	648	965	648	965
Additional Controls	Yes	Yes	Yes	Yes
Pseudo R^2	0.069	-	0.076	0.098
P-value Heteroscedasticity (Wald) Test	0.507	0.016	0.856	0.224

Note: The table reports the average marginal effects computed from different (heterosedastic) probit models with accept (columns 1 and 3) and re-follow as the dependent variables (columns 2 and 4). Columns 1 and 2 report the effects for conscientiousness while columns 3 and 4 report the effects for agreeableness/emotional stability (low levels of the respective traits as reference). Additional control variables include a dummy that equals one if the suggested user has linked a Threads account, a dummy for whether the suggested user is new to Instagram, the number of username changes for a given user, the number of emojis in the bio, a dummy for having an active story, a dummy that equals one if the user's name signals a German origin, and the follower/subscriber ratio. Robust standard errors (in parentheses) are also robust to heteroscedasticity. The Wald tests yield $\chi^2 = 0.44$ ($p = 0.507$) for model 1, $\chi^2 = 5.80$ ($p = 0.016$) for model 2, $\chi^2 = 0.03$ ($p = 0.856$) for model 3, and $\chi^2 = 1.48$ ($p = 0.224$) for model 4, indicating no significant heteroscedasticity for models 1, 3, and 4 (Heckman, 1998; Neumark, 2012) . The remaining model (model 2) is computed as heteroskedastic probit model to account for heteroscedasticity (see Section 5.5.3, p. 90). $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

significance at the five percent level and the baseline effects of higher levels of the respective personality traits increase in magnitude, the interaction effects do not systematically change the overall outcomes. Acceptance and re-follow decisions are primarily driven by the experimental treatments rather than being substantially confounded by changes in mutual connections.

Heterogeneities in Friend Suggestions Table D.1 (p. 266) presents statistically significant differences in the characteristics of potential friends suggested by Instagram, such as the number of posts, subscriptions and mutual connections. We address these heterogeneities in Tables D.7 (p. 273), D.9 (p. 275), and D.11 (p. 277) for conscientiousness and in Tables D.8 (p. 274), D.10 (p. 276), and D.12 (p. 278) for agreeableness/emotional stability by including interaction terms between the treatments (high & low agreeableness/emotional stability and high & low conscientiousness) and the respective characteristics. This allows to investigate whether the effects of our experimental manipulations vary across different subsets of suggested friends, ensuring that our primary findings are not confounded by underlying differences of suggested friends' characteristics.

Regarding conscientiousness, some of the interaction terms in Tables D.7 (p. 273), D.9 (p. 275), and D.11 (p. 277) reach statistical significance. However, most do not consistently shift the baseline effect of the treatments. In other words, a user's reaction to the friend request seems to be primarily a reaction to the treatment rather than result from differences in attributes of the suggested profiles. Although certain user characteristics significantly affected the outcome, the main treatment effects remained stable.

The same is true for agreeableness/emotional stability (see Tables D.8, p. 274, D.10, p. 276, and D.12, p. 278). The consistently positive baseline coefficient for the high agreeableness/emotional stability treatment suggests that treated users were more likely to accept or re-follow suggested profiles. While some interaction terms again exhibited a marginally or moderately statistically significant effect, they rarely overturned the main effect. Thus, whether the suggested friend had a Threads account, listed certain locations, or had additional mutual connections, tended not to substantially weaken or strengthen the difference between high versus low levels of agreeableness/emotional stability in predicting user reactions.

Concluding, individual friend-suggestion characteristics exerted only a slight moderation on the main treatment effects and did not represent a significant risk to our identification strategy.

Unobservable Heterogeneity Potential bias arising from unobservable heterogeneity between treatment groups can be mitigated by accounting for differences in error variances across these groups, which might otherwise bias standard probit estimates (Heckman, 1998; Neumark, 2012). In the present analysis, unobservable heterogeneity was identified in Model 2 of Table 5.1 and already accounted for in our main estimations. The main effects remained robust when estimating a heteroscedastic probit model instead of a conventional probit model – thus hinting at the reliability of our findings.

5.5.4 Profile Visits & Engagement

To check whether the experimental profiles were accessed by the participants and the decision to accept the friendship request or re-follow the account was not only based on the profile name and picture, we collected data on profile visits and engagement before, during, and after the experiment.[17] Figures D.2 (p. 270), D.3 (p. 270), and D.4 (p. 271) present the respective box plots on weekly profile visits, illustrating that activity levels were primarily driven by the friend requests sent during the experimental phase. Across all treatment conditions, the number of profile visits was substantially higher during the experiment compared to periods outside the experiment (see also Figure D.5, p. 271), indicating that user engagement peaked when friend requests were being sent.

[17] To that aim, profiles were created as "professional profiles" – however, without disclosing this status publicly. In addition, this setting provides access to profile statistics.

Profiles signaling low levels of conscientiousness received significantly more visits per request (2.1) than high levels (1.6), indicating a statistically significant difference at the 1% level. Profiles signaling high levels of agreeableness/emotional stability were visited more often (1.9) compared to low levels (1.6).

5.5.5 Discussion

The results of Experiment 1 show that agreeableness/emotional stability as signaled through social media cues affected the likelihood that friend requests were accepted or that users re-followed the account. Users with Instagram profiles indicating high agreeableness/emotional stability were more attractive as online friends.

Interestingly, we did not find support for high-conscientious users being preferred as online friends. The results for the re-follow rates even pointed in the opposite direction with less conscientious users being more likely to be re-followed. Potentially, conscientiousness is less important in informal settings and even less so, when there is no offline interaction to be expected. In the online world, overly conscientious people might even be perceived as a bit “boring” and threaten the user’s self-presentation if accepted as a friend (Utz et al., 2012).

Yet, we present field-based causal evidence demonstrating that visually conveying agreeableness/emotional stability on one’s social media account may substantially shape one’s ability to form a social network.

In Experiment 2, we next turn to a setting of considerably higher-stake: choosing a roommate. This allows us to test whether the agreeableness/emotional stability effect generalizes to a higher stake setting and whether conscientiousness does matter when there is more at stake and when the setting includes potential (future) offline interaction.

5.6 Experiment 2

5.6.1 Method

Experiment 2 features the four treatments included in Experiment 1 (high & low conscientiousness as well as high & low agreeableness/emotional stability) plus an additional reference treatment where our fictitious profile owner, Julia Becker, does apply for a room, but does not include a link to her social media profile.

In a between-subjects design, each room advertiser was randomly assigned to one of the five treatments, thus reducing the likelihood of detection and limiting the costs to participants. We conducted the experiment between August and September 2024 in the 15 largest student cities in Germany on the country’s largest platform for shared apartment ads. The experiment was pre-registered in the AEA RCT Registry (#12473).

In the application text for the room (which was identical across treatments), we mentioned common hobbies – such as meeting friends, jogging, watching TV series, and traveling – alongside several frequently mentioned characteristics in shared apartment applications (Moritz & Manger, 2022), including the reason for moving and previous shared-living experience.[18] To

[18] A translated version of the application text can be found in Section D.3 (p. 290).

investigate whether the application text itself conveyed any indications of the applicant's personality, we conducted a second online survey (see Section D.5.4, p. 310). The findings suggested that the text reflected moderate levels of extraversion and openness and slightly above-average agreeableness and emotional stability, as well as above-average conscientiousness.

Ads posted by professional landlords (e.g., real estate agents) were excluded as their selection behavior is likely to differ from that of potential roommates. Additionally, we excluded short-term rentals of less than six months. Further, we took measures to exclude advertisers who posted multiple ads to reduce the risk of contamination.

Apart from that, we applied to all ads that matched the criteria of our fictitious applicant in terms of age or student status and had been online for no more than three hours. To limit costs to participants, we replied to all responses within 24 to 48 hours, declining with the explanation that our fictitious applicant had already found another room at short notice.

We wrote a computer program that automatically identified suitable ads as defined above, randomly assigned treatments, and – following a review by one of the authors – sent the application to the advertiser. The program also recorded all details from each room ad, along with an extensive set of control variables, including advertiser and roommate characteristics, as well as neighborhood and district-level socio-demographics.

In total, we sent $n = 2,836$ applications to vacant room ads. We excluded 68 observations from the analysis as these stemmed from the same advertiser, availability periods were too short, the ad involved room exchanges, or the ad was posted by a rental company (which was only revealed in the reply). Thus, the final sample comprised $n = 2,768$ applications.

We collected all responses and classified them as *callback* if the response included an explicit invitation to a viewing and/or meeting with the potential roommates (or phone call). If an advertiser asked for more information, we labeled the reply as *other* response. Any explicit denial was recorded as *rejection*. As in Experiment 1, we also tracked profile visits and other engagement variables to check whether advertisers/roommates actually visited or engaged with the profiles during the experiment.

5.6.2 Results

Of the $n = 2,768$ applications, 52.9% received a response, 38.7% resulted in a callback, 5.9% were rejected, and 8.3% requested additional information about the applicant (see Table D.14, p. 282).

Figure 5.3 reports the mean callback rates across the five treatment conditions. When first comparing the difference between high vs. low levels of the manipulated personality traits, we find similar results to the ones of Experiment 1: The difference in callback rates between high and low levels of agreeableness/emotional stability (44% vs. 31%) is statistically significant ($z = -4.469$, $p = 0.000$), as the results of a two-sample Wilcoxon rank-sum (Mann–Whitney) test indicate (see also Table D.15, p. 284). This 13-percentage-point gap indicates that applicants signaling low levels of agreeableness/emotional stability via their social media profile have to send an average of 41 percent more applications to receive one callback compared to applicants signaling high levels of agreeableness/emotional stability, thus supporting H1b. In contrast and as was the case in Experiment 1, we do not find support for H2b. Although applicants signaling high levels of conscientiousness were called back more frequently compared to applicants signaling lower conscientiousness (41% vs. 37%), this difference in mean callback rates was not statistically significant ($z = -1.228$, $p = 0.219$; see Table D.15, p. 284).

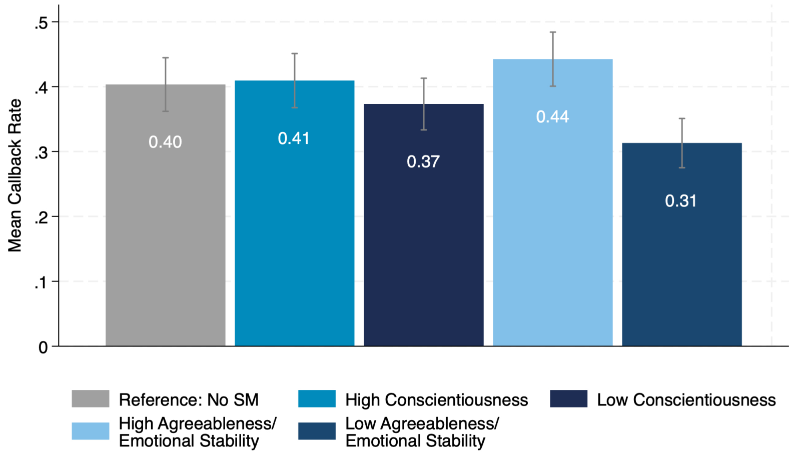


Figure 5.3: Mean Callback Rates

Regarding our second research question, we also examined the impact of providing social media information in the application for a vacant room versus not doing so (baseline treatment: No SM). Comparing the baseline callback rate of 40% to the treatments that included social media profiles, we found that low levels of the manipulated traits exhibited lower callback rates compared to the baseline, while high levels exhibited (at least slightly) higher callback rates (see Table D.16, p. 284). The difference was statistically significant only for the low agreeableness/emotional stability treatment (40% vs. 31%, $z = 3.149$, $p = 0.002$; see Table D.16, p. 284), as indicated by a two-sample test of proportion testing that the baseline and the low agreeableness/emotional stability exhibit the same callback rate, supporting H5. Differences between baseline and high agreeableness/emotional stability treatment ($z = -1.306$, $p = 0.191$), between baseline and high conscientiousness treatment ($z = -0.202$, $p = 0.840$), and between baseline and low conscientiousness treatment ($z = 1.028$, $p = 0.304$) were not statistically significant. Hence, H3, H4, and H6 were not supported.

Again, we also conducted regression analyses including various control variables. Table 5.2 presents the average marginal effects of different (heteroskedastic) probit models that estimate the effect of our treatments on the likelihood of receiving a callback. Panel A reports the results for high vs. low conscientiousness. In line with the above results, signaling high (vs. low) levels of conscientiousness had a positive, but statistically insignificant effect on the callback probability ($\beta = 0.015$, $p = 0.458$; see column 2 of Table 5.2). As displayed in Panel B, high (vs. low) levels of agreeableness/emotional stability significantly increased the callback probability ($\beta = 0.050$, $p = 0.005$). Thus, H1b is supported and H2b is not – even when including a large set of control variables on room, apartment, roommates, and other contextual variables.

Table 5.2: The Effect of Treatment Conditions on Callbacks (Average Marginal Effects)

Callback	(1)	(2)
Panel A: Conscientiousness		
Treatment: Low (Ref.)	-	-
Treatment: High	$0.0327^{\dagger}$	0.0145
	(0.0170)	(0.0196)
Pseudo R^2	0.0510	-
P-value Heteroscedasticity (Wald) Test	0.883	0.089
Panel B: Agreeableness/Emotional Stability		
Treatment: Low (Ref.)	-	-
Treatment: High	0.0583**	0.0502**
	(0.0221)	(0.0177)
Pseudo R^2	0.0520	-
P-value Heteroscedasticity (Wald) Test	0.323	0.049
Panel C: All Conditions		
Treatment: No SM (Ref.)	-	-
Low Agreeableness/Emotional Stability	-0.0752*	-0.0841**
	(0.0306)	(0.0310)
Low Conscientiousness	-0.0275	-0.0337
	(0.0312)	(0.0321)
High Agreeableness/Emotional Stability	0.0354	0.0116
	(0.0288)	(0.0264)
High Conscientiousness	0.0156	-0.00842
	(0.0259)	(0.0276)
Pseudo R^2	0.0560	0.166
P-value Heteroscedasticity (Wald) Test	0.109	0.274
Obs.	2,768	2,536
Week & City FE	Yes	Yes
Room & Shared Apartment Controls	No	Yes
Roommate Controls	No	Yes
Advertiser Controls	No	Yes
Ad Statistic Controls	No	Yes
Geographic & District-Level Demographic Controls	No	Yes

Note: The table reports the average marginal effects computed from different (heteroskedastic) probit models with callback as the dependent variable. Heteroskedastic probit models are estimated in column 2 of Panels A and B. All specifications include weekly fixed effects (dummy variables for each week during which the experiment was conducted) and city fixed effects (dummy variables for each city). The specifications in column 2 include additional control variables. Room and shared apartment controls include (among others) room size, rent, number of roommates, online time, availability, ad text characteristics, whether the ad contains pictures, and whether the advertiser offers an online viewing. Roommate controls include the sex of roommates, a dummy variable indicating if roommates speak English, dummy variables for the type of shared apartment, and other roommate characteristics. Advertiser controls include account age, a profile picture dummy, and the sex of the advertiser. Furthermore, ad statistics are controlled for by including the number of applications submitted to the respective ad and the number of responses received from the advertiser. Geographic and district-level demographic controls include distance to the city center, the number of bars in the area, a dummy variable indicating whether the apartment is located within the city, and district population metrics, such as total population, population growth, and the proportion of the female population. Standard errors (in parentheses) are clustered at the city level and are robust to heteroscedasticity. The Wald tests for Panel A yield $\chi^2 = 0.02$ ($p = 0.883$) for model 1 and $\chi^2 = 2.89$ ($p = 0.089$) for model 2, indicating no significant heteroscedasticity for model 1 (Heckman, 1998; Neumark, 2012). The Wald tests for Panel B yield $\chi^2 = 0.98$ ($p = 0.323$) for model 1 and $\chi^2 = 3.88$ ($p = 0.049$) for model 2, indicating no significant heteroscedasticity for model 1. The remaining models (column 2) of Panels A and B are computed as heteroskedastic probit models to account for heteroscedasticity. The Wald tests for Panel C yield $\chi^2 = 2.56$ ($p = 0.109$) for model 1 and $\chi^2 = 1.20$ ($p = 0.274$) for model 2, indicating no significant heteroscedasticity. $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

When comparing the social media conditions with the baseline treatment (no SM) in Panel C, we find that profiles signaling low agreeableness/emotional stability exert a statistically significant effect on the probability of receiving a callback ($\beta = -0.084$, $p = 0.007$, see column 2 of Table 5.2 including all control variables). The effect is negative. All other treatments have no statistically significant effect on the callback probability when compared to the baseline treatment. The regression results in Panel C thus confirm the above results where H5 was supported but H3, H4, and H6 were not.

5.6.3 Profile Visits & Engagement

Table D.17 (p. 285) reports the estimated average profile visits, impressions, and reach of non-subscribers per application across all treatment conditions. Overall, applications that include a social media profile signaling high levels of agreeableness/emotional stability received an average of 0.56 profile visits per application, compared to 0.48 visits signaling low levels of these traits, resulting in a statistically significant difference of $\Delta = 0.083$ ($p = 0.000$) profile visits. In contrast, average visits of profiles signaling high or low levels of conscientiousness were substantially higher with 0.73 and 0.72, respectively. This pattern was observed for all metrics. However, the differences between high and low levels of conscientiousness were substantially smaller.

Since the application text, including the profile links,[19] was equal across conditions, any observed differences in profile visits, impressions, and reach can be solely attributed to variations in the attention of advertisers, roommates, or groups of roommates. This attention may involve sharing the profile name or link among roommates to evaluate the candidate.

Interestingly, while the profile signaling low levels of agreeableness/emotional stability received the fewest visits and the lowest callback rate (see Section 5.6.2, p. 94), the profile signaling high levels of agreeableness/emotional stability had the highest callback rate, despite being significantly less visited compared to the conscientiousness conditions. Profiles portraying a conscientious individual were more effective in attracting attention but this did not increase callbacks. In contrast, the high agreeableness/emotional stability profile attracted fewer initial visits, but was more effective in converting those visits into callbacks.

5.6.4 Discussion

The results of Experiment 2 indicate that signaling high vs. low levels of agreeableness/emotional stability on one's social media profile significantly increased callbacks from potential roommates. We did not find similar results for conscientiousness. Our findings conceptually replicate the findings from Experiment 1 in a higher-stake setting, thus highlighting the role of visual personality cues for interpersonal relations.

When comparing the callback rates with the baseline treatment where no social media information was provided, we found the effects of high vs. low agreeableness/emotional stability to be driven by the deterrent effect of the low agreeableness/emotional stability treatment rather than the appeal of the high agreeableness/emotional stability treatment.

Interestingly, profiles signaling high or low levels of conscientiousness resulted in increased profile visits and engagement compared to profiles signaling different levels of agreeableness/emotional stability. The reason could be that conscientiousness is usually not the primary trait that social media users express on social media; they usually want to appear attractive, extraverted,

[19] The profile names used are juli1becker, _juli_becker_, _juli1becker_, and _juli_becker1, which are highly similar and unlikely to cause differences in clicks.

and popular (Scott, 2014), so people might have checked out these profiles out of curiosity. Profile visits did not automatically translate into callbacks, indicating that participants processed the information provided. Profiles signaling high agreeableness/emotional stability were most successful in converting visits into callbacks.

Prospective roommates seem to be reluctant to share their apartment with a person who displays low levels of agreeableness/emotional stability. Whether an applicant displays a high level of agreeableness/emotional stability in their attached social media profile or does not provide any social media information, seems to be of considerable less importance – at least in a situation where the application text already indicates slightly above average levels of agreeableness and emotional stability.

5.7 General Discussion

The aim of this paper was to examine whether visual personality cues in social media profiles do not only affect impression ratings, but also have effects on real-world outcomes. The results of two large-scale randomized field experiments showed that people with profiles signaling high agreeableness/emotional stability were more successful than people with profiles signaling low agreeableness/emotional stability: they were more likely to be accepted as online friends and as potential roommates. Interestingly, differences in conscientiousness (high vs. low levels) did not affect the outcomes in both settings.

Our randomized field experiments make several important contributions. First, they demonstrate the power of social media-based cues in shaping real-world outcomes: We found that the profile signaling high levels of agreeableness/emotional stability exhibited higher acceptance and callback rates than the profile signaling low levels of agreeableness/emotional stability. We go, thus, beyond prior work demonstrating that information on fictitious social media profiles viewed during an (online) experiment affects personality judgments (Utz, 2010; Walther et al., 2009; Walther et al., 2008). Second, Experiment 2 sheds additional light on the underlying processes because here we could also include a baseline treatment where no social media information was provided. The comparison with the baseline treatment showed that people especially avoided targets whose profiles signaled low agreeableness/emotional stability. This makes sense because relationships with these are known to be characterized by a higher level of conflict (Jensen-Campbell et al., 2002). Detecting and avoiding these people also makes sense from an evolutionary perspective. Selecting the right interaction partner was linked to survival and reproduction success. The ability to select people with whom one has less conflicts might, thus, have evolved in humans (Buss, 1991). Our work demonstrates that this detection also works with subtle signals detected by machine learning algorithms such as bright or hues of colors (Cucurull et al., 2018; Segalin et al., 2017) and not only with relatively obvious cues like many photos displaying the target with friends.

An interesting finding in itself is that we were not able to independently manipulate agreeableness. In two randomized pilot experiments, manipulations of agreeableness also affected the assessment of other personality traits. While we were successful in reducing the unintended effects on perceived extraversion and openness, the emotional stability ratings remained affected by the agreeableness manipulation. An explanation could be that we manipulated signals identified in papers using machine learning (Cucurull et al., 2018; Segalin et al., 2017). Maybe people are not as good as algorithms to detect very subtle differences. Another explanation could be that the emotional stability ratings are affected by a spill-over effect. Agreeableness has been identified as a very central trait in a recent meta-analysis (Wilmot & Ones, 2022). Judgments on this very central dimensions might automatically color judgments on emotional stability as another trait highly relevant for interpersonal interactions, especially conflicts (Harris & Vazire, 2016). Future research is needed to disentangle the two personality traits.

Our results also contribute to work on the importance of conscientiousness. In contrast to our hypotheses, we did not find effects of conscientiousness on the acceptance of friend requests or callback rates. This is puzzling because the conscientiousness manipulation worked well in both pilot experiments. Descriptively, participants were even more likely to re-follow the people whose social media profiles signaled low rather than high conscientiousness. One could argue that, in the online world, overly conscientious people might be perceived as a bit "boring," and users might want to avoid them being displayed as their social media friends – with prior research having shown that the perceived personality of social media friends also affects personality ratings of the profile owner (Utz, 2010; Walther et al., 2008) and that social media users want to appear popular (Utz et al., 2012).

Appearing popular would, however, not seem to be a prominent motive when selecting potential roommates. When sharing an apartment, it is important whether roommates stick to the cleaning schedule, for example. However, while conscientious people are generally preferred as roommates according to our first online survey, we did not find any effects of conscientiousness in the corresponding field experiment (Experiment 2). We used a student sample, though, so conscientiousness in other areas, such as studying hard, not drinking or exercising regularly might also be perceived as less attractive, depending on who already lives in the shared apartment. It could also be the case that scoring high on agreeableness/emotional stability is important for the decision on whom to invite for a room tour and that a (sufficiently high) level of conscientiousness is then secured at later stages of the selection process. Another explanation for not having found an effect for conscientiousness in Experiment 2 is that the application text explicitly stated that the target is familiar with cleaning schedules; this might have reduced a potentially negative effect of displaying low levels of conscientiousness. Follow-up field research could examine whether conscientiousness matters more in other contexts, e.g., recruiting.

Further, we improved prior experimental studies that used fictitious profiles in hypothetical settings by analyzing not only real-world decisions, but also analyzing additional data on profile visits and engagement. We found that the fictitious profiles and their associated information were actively accessed and evaluated by potential friends and roommates, as evidenced by significant increases in engagement metrics during the experimental periods. Higher levels of profile visits and engagement did, however, not consistently translate into increased acceptance, re-follow, or callback rates. For instance, in Experiment 2, the profile signaling high agreeableness/emotional stability received most callbacks on average, although it was less visited than the profiles signaling conscientiousness. This indicates that people process the information they encounter on profiles, and that attention and subsequent behavior are triggered by different processes. The low/high conscientious profiles correspond less to the stereotypical social media presentations and might, thus, receive more visits. It could also be that the person who initially screened the applications more eagerly shared the corresponding links with other incumbent roommates, maybe especially for more difficult to decide cases. The pattern here is another hint for the claim that the importance of conscientiousness for the selection of a roommate is less clear. If people clicked on a highly agreeable profile, they more often decided to call the target back, confirming prior work on the importance of agreeableness (Wilmot & Ones, 2022).

Before closing, we would like to note some strengths and limitations of our paper. A clear strength is that we examined the role of profile characteristics in real-world settings. A second strength is the scale of the experiments – we sent roughly 1,000 friend requests and applied to about 2,800 vacant rooms. Third, we replicated the finding that high vs. low agreeableness/emotional stability cues mattered in a second considerably higher-stake setting. Fourth, by manipulating personality mainly by pictures (the text in the captions was only visible once people clicked on the pictures), we bridged the classical lexical approach to personality (Allport & Odbert, 1936; McCrae & Costa, 1987) with an image-based approach to signal personality (Cucurull et al., 2018; Segalin et al., 2017) – assessing the causal impact of visual personality cues on real-world decisions.

One limitation of our paper is that we did not succeed in manipulating perceived agreeableness without also affecting the judgments of emotional stability. A second limitation is that we focused on young adults, mainly students. This group grew up with social media; the effects might be different for other populations.

Future field studies might seek to explore additional personality traits, differences in gender, or examine whether conscientiousness becomes more salient in professional or performance-based contexts. The causal evidence from both field experiments underscores the power of visual cues signaling agreeableness/emotional stability and how digital self-presentation and impression formation from social networks affect on- and offline dynamics – an omnipresent and pervasive phenomenon in an increasingly interconnected world.

Chapter 6

ESG Incentives in Executive Compensation[1]

"ESG metrics are now one of the most prevalent metrics in executive incentive plans"
WILLIS TOWERS WATSON (2023)

6.1 Introduction

In recent years, there has been a growing trend of incorporating Environmental, Social, and Governance (ESG) metrics into executives' performance-linked compensation plans, as observed by many industry experts and scholars (e.g., COHEN, KADACH, ORMAZABAL, & REICHELSTEIN, 2023; HAZARIKA et al., 2022; PRICEWATERHOUSECOOPERS, 2021; WILLIS TOWERS WATSON, 2023). Typical ESG metrics include, for instance, emissions, working conditions, workforce diversity, or employee satisfaction. This shift in ESG contracting raises important questions: What motivates firms to link executive compensation to ESG outcomes, and does this practice genuinely create incentives for executives to improve ESG performance?

To measure the adoption of ESG-linked pay, past research has primarily focused on identifying whether a given company incorporates at least one ESG performance metric (ESG criterion) in the compensation contract of at least one of its executives.[2] The information necessary for this approach is readily available and allows the analysis of large (global) firm samples.[3] However, this approach does not reveal the actual importance that firms assign to ESG metrics. For example, it cannot refute the possibility that firms report ESG metrics to "greenwash" executive pay: i.e., firms could be windowdressing incentive pay to appear ESG-friendly but without making

[1]Chapter 6 is based on the working paper "All Hat and No Cattle? ESG Incentives in Executive Compensation" by Matthias Efing (HEC Paris), Stefanie Ehmann (University of Tübingen), Patrick Kampkötter (University of Tübingen), and Raphael Moritz (University of Tübingen). The HEC Paris Research Paper No. FIN-2024-1506 is available on SSRN: http://dx.doi.org/10.2139/ssrn.4974204.

[2]See, for example, TSANG et al. (2021), HAZARIKA et al. (2022), QIN and YANG (2022), ARESU et al. (2023), CARTER et al. (2023), COHEN, KADACH, ORMAZABAL, and REICHELSTEIN (2023), IKRAM et al. (2023), and BARONTINI and HILL (2024).

[3]As such, HAZARIKA et al. (2022) and COHEN, KADACH, ORMAZABAL, and REICHELSTEIN (2023) are able to study, for example, regulatory and cultural differences across countries as determinants of ESG adoption in executive pay.

ESG targets sufficiently ambitious or without giving them material weight in the remuneration function.

In this paper, we move beyond this ”extensive margin approach” to measuring ESG adoption in executive pay. Instead, we use a novel dataset with comprehensive information both on the ex-ante design of executives' compensation plans (such as numbers and weights of different performance metrics or target bonuses) and on realized performance (e.g., target achievement rates) and payout (e.g., salary, bonuses, etc.). This detailed analysis results in a smaller firm sample than in previous research; still we are able to analyze a panel of 674 executives from 73 constituent firms in Europe's leading stock indices, the EURO STOXX 50 and the STOXX Europe 50, between 2013 and 2020. Crucially, the joint analysis of ex-ante contract design and realized pay helps gauge firms' main rationales for linking compensation to ESG.

A priori, there are several possible reasons why firms would tie executive compensation to ESG criteria. For example, strengthening ESG performance could increase shareholder value (doing well by doing good). Firms may then condition executives' incentive pay on ESG performance metrics, especially if strong ESG performance acts as a leading indicator of shareholder value (e.g., Cohen, Kadach, Ormazabal, & Reichelstein, 2023; Edmans, 2023; Edmans et al., 2017).[4] When stronger ESG performance does not increase, but undermines shareholder value, the rationale for ESG-linked pay is more complex. Still, incentivizing executives to improve ESG performance can be efficient if the firm's objective is not pure shareholder value maximization but places some weight on ESG, for example, because investors genuinely care about the environment and society (shareholder welfare approach).[5]

However, ESG-linked pay of executives can also occur for reasons other than incentive provision. Firms can “greenwash” executive pay to appease external pressure by third parties without sacrificing shareholder value to ESG performance (Cho & Roberts, 2010; Christensen et al., 2018; Crilly et al., 2016; Grabner et al., 2024). For example, firms might communicate publicly about a large number of ESG metrics in executives' incentive pay without giving these measures a significant weight in the calculation of realized bonuses and other pay elements. Alternatively, they can design ESG metrics in a way that achievement is certain, making ESG-linked performance pay essentially part of (guaranteed) fixed salary with little incentive power. Finally, entrenched executives themselves could seek the inclusion of easy-to-achieve or hard-to-verify ESG performance metrics as a way to extract rents from the firm (e.g., Bebchuk & Tallarita, 2022).[6] Under such greenwashing and rent-seeking, ESG-linked compensation does not generate additional ESG incentives for executives.

To find out which rationals explain best the rise of ESG-linked pay, past research has tried different approaches. For example, existing papers have studied whether firms that include ESG metrics in executive pay show higher ESG performance or how such ESG adoption correlates with different investor clienteles, industries, regulatory regimes, or cultural differences across

[4]If markets do not fully understand the effect of ESG on long-term value creation, ESG outcomes convey information not already included in the stock price. The informativeness principle by Holmström (1979) implies that firms would then condition compensation also on ESG outcomes (rather than on the stock price alone).

[5]See, for example, Broccardo et al. (2022), Hart and Zingales (2022), Landier and Lovo (2024), Barzuza et al. (2023), and Bonham and Riggs-Cragun (2024).

[6]For a discussion of how tax incentives or ESG subsidies might impact greenwashing behavior or managers' behavior to game the ESG-related incentive scheme, see, for instance, Bonham and Riggs-Cragun (2024).

countries (e.g., Aresu et al., 2023; Carter et al., 2023; Cohen, Kadach, Ormazabal, & Reichelstein, 2023; Hazarika et al., 2022).[7] This literature generates important insights, which we review further below. However, we believe that it lacks a natural first step: a detailed analysis of the *design* of ESG-linked compensation. For example, the simple weight of an ESG metric in the calculation of performance pay can already tell a lot about whether this metric has any hope of altering incentives. Similarly, a variance decomposition of realized performance pay may tell us whether executives' wealth, and thus their financial incentives, depends mostly on variation in ESG or non-ESG performance.

In a first step, our own analysis adopts the commonly used extensive margin approach to quantify how much firms include ESG performance metrics in executive pay. In our European sample, this method suggests that ESG is indeed becoming more significant, with the share of executives with short-term incentive pay (STI) tied to at least one ESG criterion increasing from 40% in 2013 to 60% in 2020.[8] Social metrics related to the workforce and product quality/responsibility are particularly widespread, while environmental metrics, and social metrics related to local communities and human rights, remain less common. For example, emissions-linked performance metrics only become more prominent after 2018, coinciding with a fourfold increase in the price of EU carbon permits. Still, the 60% share of executives with at least one ESG metric seemingly confirms that the importance of ESG performance metrics in executives' incentive pay is growing.

We then move beyond the extensive margin approach and study the design of compensation plans in more detail. Our granular data reveal that the role of ESG in STI is in fact limited. First, ESG metrics are far less prevalent than non-ESG metrics, with approximately four times fewer ESG metrics. Moreover, many of these ESG metrics are discretionary, meaning that they are assessed jointly with non-ESG metrics at year-end. In other words, for these discretionary ESG measures, the supervisory board or compensation committee determines their exact weight in the calculation of performance pay only at the end of the fiscal year, leaving the executive uncertain about how much ESG performance will, in fact, be rewarded. Even when ESG metrics are binding, i.e., when the firm commits to weights for ESG metrics at the beginning of the year, they enter the calculation of realized STI pay only with a small weight of 2% in 2013 and 5% in 2020 for the average executive in our sample. This is arguably too small for a material effect on executives' incentives.

After studying the ex-ante design of compensation plans, we turn to ex-post pay realizations, and ask how much variation in STI is driven by variation in ESG or non-ESG performance. In a subsample of 57 firms that report granular target achievement rates for different metrics, we decompose the variance of realized STI into different components for ESG and non-ESG

[7] Similar studies can be found on the adoption of CSR criteria in executive compensation (e.g., Al-Shaer et al., 2023; Flammer et al., 2019; Hong et al., 2016; Ikram et al., 2023).

[8] Whenever available, we hand collect and report the necessary information both for short-term incentive pay (STI) with a performance period of one year and for long-term incentive pay (LTI) with longer performance periods. For some parameters such as target achievement rates, reporting granularity and data availability is better for STI plans. For instance, LTI target values are reported in a minority of cases, with information often missing in earlier years. Additionally, LTI details are sometimes not specified at the individual level, but only for the entire executive team. Lastly, it often occurs that multiple LTI plans for the same executive overlap across time, as each of these plans have different vesting periods, making it difficult to determine LTI target achievement. Nevertheless, results regarding the adoption of ESG criteria in STI and LTI plans are qualitatively similar.

performance pay. Binding *non-ESG* metrics (binding metrics have an explicit weight to which the firm commits ex ante) account for 87%, that is, for the lion's share of STI variance. Binding *ESG* metrics account only for an immaterial 1%. The remaining 12% of STI variance are explained by 8% for discretionary metrics (ESG and non-ESG metrics that are assessed jointly without individual ex-ante weights), and 4% for covariance terms.[9] Even if we drop all STI plans without any ESG metric and consider only the remaining subsample of ESG adopters, binding ESG metrics contribute only 2.5% to total STI variance, that is, 26 times less than binding non-ESG metrics.

Of course, the small contribution of ESG metrics to STI variance is largely due to the small weight that firms assign to ESG in the payout function. However, it is interesting to note that the variance share of 1% for binding ESG metrics even falls short of their official weight, which implies that target achievement is in fact more stable for ESG than for non-ESG metrics. Additional analyses confirm that executives with more ESG metrics and higher ESG target weights are exposed to less pay risk overall, as their overall target achievement for all targets combined is more stable. It seems that ESG metrics are either less ambitious than non-ESG metrics or that ESG performance measurement is less noisy.[10]

Overall, our analyses of ex ante contract design and ex post STI pay realizations reveal an important dichotomy. On the one hand, the weight of ESG in the payout function and, hence, the contribution of ESG to pay risk are very small, which makes it difficult to explain ESG-linked pay under rationals of incentive contracting or rent extraction. On the other hand, the share of executives with at least one ESG metric increases, and so does the number of ESG metrics per executive. This begs the question of why firms report an increasing number of ESG metrics, albeit ESG metrics with limited hope of altering incentives? To explore this further, we analyze heterogeneity in executive pay across industries, firms, and executive positions, which yields several key insights.

First, there are in fact some industries where ESG-linked pay appears to align more closely with the rational of incentive contracting. Specifically, sectors with a significant environmental impact, such as energy and utilities, stand out. These industries are particularly vulnerable to rising energy costs and EU carbon permit prices, potentially making policies that reduce emissions or conserve resources not only environmentally beneficial but also financially advantageous.

In such industries, executive pay is indeed linked to more binding ESG metrics with larger weights, and, hence, ESG also explains a larger share of STI variance. Interestingly, our variance decomposition shows that the covariance between ESG and non-ESG target achievement is positive and six to seven times higher in the energy and utilities sector than for the average executive. This is consistent with the hypothesis that ESG performance and financial performance are positively associated in the energy & utilities sector (doing well by doing good). Importantly, executive pay in energy-dependent and emissions-heavy industries is rarely linked to discretionary ESG metrics, that is, to performance measures that firms can choose to ignore

[9] These numbers are reported for a variance contribution that only considers within-executive variation in realized STI. We also decompose total (within- as well as between-executive) variation, in which case achievement rates of binding ESG targets explain 2% of total STI variance.

[10] A third explanation could be that combining ESG and non-ESG metrics reduces overall pay risk through diversification. However, the covariance term between ESG and non-ESG target achievement is relatively small.

at year-end. In contrast, such discretionary ESG metrics are prevalent in the financial services sector, where binding ESG metrics (with arguably larger incentive power) are largely absent.

These findings of our industry analysis also hold within industry, i.e., in regressions with industry fixed effects: Firms with a historically high carbon footprint avoid discretionary ESG metrics, while large, visible companies subject to scrutiny from institutional (but dispersed) investors and independent directors are more likely to include numerous discretionary ESG metrics. These same companies often fail to link executive pay to binding ESG metrics with material weights, weakening the potential of ESG criteria to change incentives. By contrast, firms with highly volatile stock prices – where stock performance is a less reliable measure of individual performance – are more likely to implement binding ESG metrics with substantial weights in STI calculations. This aligns with the prediction that when stock prices provide a poor signal of executive actions, firms compensate by incorporating additional metrics, such as ESG, into incentive pay.

Overall, our analysis of different industries and firms suggests that there is no single rationale for ESG-linked pay. Sure, for the vast majority of companies, ESG plays a minimal role in executive compensation. Especially financial firms and large, visible companies under public scrutiny tend to rely on discretionary ESG metrics with arguably limited incentive power, which could be driven by greenwashing strategies. However, for a subset of firms, particularly in energy-intensive and high-polluting sectors, as well as firms with highly volatile stock prices, (binding) ESG criteria could play a more meaningful role within an incentive contracting framework.

Lastly, we compare compensation plans across various executive positions to determine whether firms tailor ESG performance metrics to the specific responsibilities of different top managers. Intuitively, one would expect workforce-related metrics to be more prominent for a Chief Human Resource Officer (CHRO), or environmental and emissions-related metrics to be assigned to a Chief Technology Officer (CTO). Surprisingly, our findings suggest otherwise. In firm-year fixed effects regressions, CHROs are no more likely to have workforce-related metrics than the CEOs of the same company, and CTOs are actually significantly less likely to have environmental metrics than the corresponding CEOs in the same firm-years. This is unexpected from an incentive contracting perspective, where one would anticipate a stronger alignment between specialized executives, such as the CHRO or CTO, and their respective ESG-related tasks. The CEO, as a generalist responsible for overall firm performance, could reasonably have pay tied primarily to stock price, rather than specific ESG outcomes. More broadly, we find that ESG-linked pay is generally less common among specialized C-suite roles and most frequently applied to more visible generalist positions, particularly CEOs. This pattern seems suggestive of firms applying ESG metrics to their most publicly visible executives, potentially as part of a greenwashing strategy.

The rest of the chapter is organized into seven sections. We review the literature on ESG (and CSR) contracting in Section 6.2. In Section 6.3, we develop testable hypotheses regarding ESG-linked executive pay under incentive contracting, greenwashing, and rent extraction. The typical design of STI and LTI contracts in practice is described in Section 6.4. Section 6.5 describes the data and Section 6.6 our empirical analyses. We conclude in Section 6.7.

6.2 Literature Review

ESG (or CSR) contracting describes the integration of non-financial sustainability and corporate social responsibility metrics in executive pay contracts (e.g., Kolk & Perego, 2014). These terms all describe “a firm’s voluntary actions to manage its environmental and social impact and increase its positive contribution to society.” (Khan et al., 2016, p. 1697).[11] More recent contributions find that ESG/CSR contracting is associated with firm-level outcomes such as long-term orientation, financial and social performance outcomes, and improvements in environmental and social initiatives (e.g., Al-Shaer et al., 2023; Flammer et al., 2019; Hong et al., 2016; Ikram et al., 2023). It has also been shown that particularly quantitative ESG performance measures can effectively generate incentives to improve ESG outcomes (Maas, 2018).[12] However, this evidence might be subject to limited generalizability, as the vast majority of these studies uses US data (e.g., Flammer et al., 2019; Grabner et al., 2024; Ikram et al., 2023) and ESG contracting has been shown to be less developed in the US than, for instance, in Europe (Barontini & Hill, 2024; Cohen, Kadach, Ormazabal, & Reichelstein, 2023; Hazarika et al., 2022).

Most recently, contributions have shifted their attention towards a more international perspective of ESG contracting, documenting the prevalence of ESG metrics in executive pay plans and their anticipated effects. Cohen, Kadach, Ormazabal, and Reichelstein (2023) analyze a broad sample of 4,395 listed firms from 21 countries between 2011 and 2020 and find that firms in more resource-intensive sectors as well as those with higher engagement by institutional investors are more likely to engage in ESG contracting. Ultimately, ESG adopters report lower CO2 emissions and attract higher ESG scores from rating agencies. Barontini and Hill (2024) analyze a sample of 53,602 listed firm-year observations from 58 countries and 19 industrial sectors over the period 2002-2021 and broadly confirm previous results on adopters of ESG contracting, based on a dummy variable which reflects whether a firm has implemented an ESG- or sustainability-oriented executive compensation policy. Hazarika et al. (2022) expand the analysis of Cohen, Kadach, Ormazabal, and Reichelstein (2023) and include a country’s culture as main predictor of ESG pay adoption. In their analysis of firms from 59 countries in the time period 2005-2020 they furthermore exploit a regulatory shock that requires firms to disclose ESG-related information and find that ESG contracting leads to higher social and financial outcomes, most likely mediated by employee satisfaction. Carter et al. (2023) exploit the staggered adoption of say-on-pay (SOP) voting laws in 36 countries between 2002 and 2019 to show that the prevalence of ESG contracting increases, particularly in jurisdictions with binding SOP votes.

The influence of institutional characteristics such as regulatory pressure on ESG contracting is also researched by Aresu et al. (2023), who exploit a dataset of 2,328 firms listed in 37

[11] ESG and CSR contracting have been used interchangeably so far. The term ESG is now used to group all sustainability-related issues (Khan et al., 2016). More recently, Edmans (2024) has added the term “rational sustainability” to this debate.

[12] Moreover, the role of corporate governance aspects in ESG contracting is of growing interest to researchers. Recent papers emphasize, for instance, the pivotal role of sustainability committees in this context and the complementary relationship between the public disclosure of sustainability information and CSR contracting (Al-Shaer & Zaman, 2019; Grabner et al., 2024).

countries in 2003 through 2015. Using a firm's first-time adoption of CSR contracting as the dependent variable, they find that firms in countries with greater social and environmental regulatory pressures are more likely to adopt ESG contracting, but that this effect is moderated by a firm's internal corporate governance structure such as block ownership or board independence. Tsang et al. (2021) add the moderating role of institutional settings on the relationship between CSR contracting and firm innovation to the discussion, based on a sample of 17,855 firm-year observations from 30 countries between 2004 and 2015.

The empirical papers cited above mostly measure ESG adoption in executive pay with a binary variable that equals one if a firm adopts at least one ESG element or a broader ESG-related compensation policy, and thus can make only limited statements regarding the importance that such ESG adoption takes relative to non-ESG criteria.[13] Not surprisingly, researchers explicitly have called for more detailed research on ex-ante information in pay contracts: *"Specifically, it would be valuable for future research to have further access to the exact compensation vehicles, the relative weights attached to different performance metrics, and the use of discretionary bonus rules* (Cohen, Kadach, Ormazabal, & Reichelstein, 2023, p. 810)." We contribute to the literature a detailed analysis of the ex-ante design of compensation plans. In particular, we distinguish between binding performance metrics (whose weights are fixed by the firm ex ante) and discretionary metrics, and compare the relative weights that firms assign to ESG and non-ESG criteria. As regards ex-post realizations, we collect individual target achievement rates, which allows us to identify the relative contribution of ESG and non-ESG target achievement to the overall pay risk that performance-linked remuneration poses to executives. Importantly, we uncover a number of novel insights regarding heterogeneity in contract design across industries and executive positions.

Finally, ESG-linked executive pay has also recently been studied in the theoretical literature. For example, Bonham and Riggs-Cragun (2024) analyze a principal-agent model where boards contract with managers to maximize shareholder welfare, denoted as a sum of financial and ESG outcomes. In a scenario where shareholders treat financial and ESG outcomes as complements, desired behavior of the agent can be triggered with stock awards that are contingent on ESG performance, or, in other words, incentives for financial performance are increasing in ESG performance.

6.3 Hypothesis Development

In this section, we discuss different rationales for why firms include ESG performance metrics in executive incentive pay and develop several testable hypotheses. The first rationale we consider is that (some dimensions of) ESG performance could increase shareholder value (doing well by doing good). For example, higher employee satisfaction (a common ESG criterion) could result in a higher stock price if it leads to higher motivation and more successful recruitment (e.g., Edmans, 2011). Hence, some firms might include employee satisfaction as a key performance

[13] Cohen, Kadach, Ormazabal, and Reichelstein (2023) at least document median weights of ESG metrics in STI and LTI bonus plans, but do not report further analyses based on these weights. Bebchuk and Tallarita (2022) document in a cross-section of 2021 S&P 100 firms that only less than one-third of firms report ex-ante weights of ESG metrics.

indicator in executive compensation. However, this begs the question of why these firms would not link compensation directly to the stock price. If the ultimate goal is shareholder value creation, would it not be more efficient to use *pure* equity incentives and to let the executive decide whether she should allocate her time and effort to employee satisfaction or to some other determinant of shareholder value?

One common argument against pure equity incentives derives from the informativeness principle by Holmström (1979), which states that the principal optimally conditions compensation on additional signals if these signals provide new information about the agent's past actions.[14] In our context, ESG outcomes could provide information that is not contained in the stock price for several reasons. For example, markets might not be fully efficient and fail to understand the effect of ESG outcomes on long-term value creation, such that ESG outcomes act as leading indicators of the stock price (e.g., Cohen, Kadach, Ormazabal, & Reichelstein, 2023; Edmans et al., 2017).[15] Furthermore, if the stock price is very noisy and depends on factors outside the executive's control, it might be efficient to condition compensation on non-equity criteria with a closer link to the executive's job (possibly ESG criteria). For example, the link between a CHRO's actions and employee satisfaction is likely stronger (and easier to understand) than the link between the CHRO's actions and the stock price, which depends on firm-wide performance and thus on corporate policies she cannot control. In general, the more specialized the job of a given executive, the less informative should be firm-wide performance, as captured by the stock price, about the executive's actions. In that case, ESG performance metrics tailored to the job of the specialist could have a relatively higher "signal-to-noise ratio". By contrast, equity incentives should be more appropriate for a generalist like a CEO, whose job is to set the overarching, long-term strategy and vision for the *entire* firm and to integrate all different corporate policies in an attempt to maximize firm-wide performance. Overall, the noisier the equity (as a measure of the executive's actions) the more relevant could become ESG performance metrics for executive pay.

Our example of a positive link between employee satisfaction and financial performance falls under the paradigm of "doing well by doing good". However, firms sometimes face trade-offs between (some dimensions of) ESG performance and shareholder value. For example, a firm might find it *financially* optimal to ignore concerns about workplace conditions further up its supply chain rather than to perform costly due diligence in its suppliers' factories. This might be especially true if bad workplace conditions up the supply chain are difficult to detect for third parties and the risk of reputational damage in the future is therefore low. For such a trade-off between financial and ESG performance, the objective of shareholder value maximization implies that the firm does not tie executive pay to ESG metrics in any meaningful way that affects incentives. However, this does not preclude the possibility that the firm pretends to be "ESG conscious" and, for example, includes non-consequential ESG criteria in executive pay. For example, firms can artificially report a large number of ESG performance metrics in executive pay but without giving these measures a material weight in the payout function. Firms can

[14] Also see Lambert and Larcker (1987).

[15] Krueger et al. (2020) show that many institutional, long-term oriented investors believe that climate risk has financial implications that are not fully priced into firms' equity, and Garel et al. (2024) show that a biodiversity risk premium emerged only after the Kunming Declaration in 2021.

also tie compensation to ESG metrics that the executive is sure to achieve, such that "ESG performance pay" essentially becomes part of fixed salary without any incentive effects.

Firms might engage in such "greenwashing" of executive pay if meaningful (that is effective) ESG incentives would conflict with shareholder value maximization but, at the same time, firms face public pressure by third parties (governments, customers, some investor groups, proxy advisors, etc.) to become more ESG-friendly. Although it might not be possible to deceive (rational) third parties in equilibrium over multiple periods, firms could still engage in greenwashing executive pay during some transition phase of rising ESG pressure, especially if greenwashing is difficult to detect in the short run and cheap to implement (Cohen, Kadach, Ormazabal, and Reichelstein 2023, p. 812).

If we relax the assumption of pure shareholder value maximization and consider the possibility that sufficiently many shareholders genuinely care about ESG outcomes (e.g., Landier & Lovo, 2024) and wish to maximize shareholder *welfare* rather than shareholder *value* (e.g., Bonham & Riggs-Cragun, 2024; Hart & Zingales, 2022), then firms will optimally tie executive pay to ESG outcomes even if doing so sacrifices some shareholder value. These ESG metrics would enter executive pay with a material weight, and their achievement would not be automatic but truly depend on the actions chosen by the manager.

Summarizing the above discussion, we expect **incentive contracting** on ESG outcomes in two scenarios: (1) when better ESG performance increases shareholder value and pure equity incentives would be a too noisy signal of managers' actions; and (2) when better ESG performance decreases shareholder value, but shareholders genuinely care about ESG and are willing to sacrifice shareholder value for better ESG outcomes. In contrast, under **greenwashing**, ESG metrics in executive pay are immaterial and/or target achievement is de facto guaranteed such that ESG pay is effectively part of fixed pay without any incentive effects, while still seemingly responding to potential investor pressure.

Overall, we make the following predictions for our empirical analysis:

HYPOTHESIS 1. *Relevance of ESG Metrics*

(a) ***Incentive contracting:*** *ESG metrics have a material weight in the calculation of executives' performance pay. ESG target achievement is not guaranteed.*

(b) ***Greenwashing:*** *Firms report ESG metrics but do not assign them any material weights in the calculation of performance pay, and/or target achievement is quasi-automatic, suggesting that firms simply relabel part of executives' fixed salary as ESG performance pay.*

HYPOTHESIS 2. *Industry and Firm Characteristics*

(a) ***Incentive contracting:*** *ESG metrics with material weights and non-automatic achievement are more common in industries whose shareholder value increases in ESG performance and in firms whose stock price is a relatively noisy measure of executives' actions.*

(b) ***Greenwashing:*** *ESG metrics with immaterial weights and/or automatic achievement are more common in industries without a strong positive relation between shareholder value and ESG performance and in firms that attract more public attention and ESG pressure.*

HYPOTHESIS 3. *Executive Characteristics*

(a) ***Incentive contracting:*** *ESG metrics are tailored to the job of the executive. They are more common among specialized executives responsible for only a limited number of corporate policies or functions than among CEOs and other generalists in charge of firm-wide performance and strategic goals.*

(b) ***Greenwashing:*** *ESG metrics are unrelated to the specific task performed by the executive, and more common among more visible executives under public scrutiny. Specifically, ESG metrics are more common among CEOs than among lower-ranked, specialized executives.*

Finally, a third rationale for ESG pay, apart from incentive contracting and greenwashing, is rent extraction. The general prediction is that entrenched executives convince shareholders/directors to tie their compensation to dimensions of performance on which they expect to succeed (Morse et al., 2011, Edmans et al. 2017, p. 467, Bebchuk and Tallarita 2022).

We note that a large number of reported ESG metrics albeit with *immaterial* weights can occur under greenwashing but not under rent extraction, as entrenched executives want their preferred performance metrics to have a *large* weight in the calculation of their realized performance pay. We can also disentangle rent extraction from incentive contracting because rent extraction predicts that executives choose ESG metrics that they are sure to achieve and thus expose them to very little pay risk. By contrast, incentive contracting is inconsistent with such guaranteed target achievement.

HYPOTHESIS 4. *Rent Extraction*
Entrenched executives convince shareholders and supervisory boards to tie their performance pay to ESG metrics they are sure to achieve. These performance metrics enter the calculation of performance pay with a material weight, but target achievement is quasi-automatic, making ESG pay essentially part of fixed salary.

6.4 Contract Design in Practice

An executive's compensation contract is typically composed of base salary, annual variable compensation, and multi-year variable compensation elements (see, for instance, Murphy (1999) and Edmans et al. (2017) for a detailed description of the structure of executive pay and Edmans et al. (2023) for a survey on the objectives, constraints, and determinants of CEO pay from both a director's and investor's perspective).

The annual base salary (also referred to as fixed compensation) includes all contractually agreed fixed compensation components, which are paid regardless of individual, business, or firm-wide performance. In the following, we describe in more detail executives' variable (i.e., performance-contingent) remuneration.

6.4.1 Short-Term Incentive Pay (STI)

Executives' *annual* variable compensation, denoted as short-term incentive pay (STI), includes all variable compensation elements with a performance period of one year. Its calculation is generally formula-based.[16] At the beginning of the year, firms set a *target* bonus or *target* STI amount, which we denote as $TSTI_{i,t}$ for executive i in year t. The *realized* bonus or *realized* STI amount $RSTI_{i,t}$ at the end of the year depends on the extent to which the executive has achieved the target, i.e., on the fulfillment rate $f_{i,t}$:[17]

$$RSTI_{i,t} = TSTI_{i,t} \times f_{i,t} \tag{6.1}$$

Usually, the calculation of $RSTI_{i,t}$ in Equation 6.1 is qualified by so-called bonus hurdles and bonus caps. Specifically, firms impose constraints on the fulfillment rate $f_{i,t}$. If actual performance falls below a lower threshold (the hurdle), realized $RSTI_{i,t}$ is set to a lower bound (typically zero). Similarly, $RSTI_{i,t}$ is sometimes capped at an upper threshold (the cap). In the so-called incentive zone between the hurdle and the cap, $RSTI_{i,t}$ increases (typically linearly) in target fulfillment $f_{i,t}$ as in Equation 6.1.

In practice, most contracts include several performance metrics. It then becomes important how much each of these different metrics contributes to overall target achievement $f_{i,t}$. Most bonus plans (STI plans) in our sample are additive:

$$f_{i,t} = \underbrace{w_{i,t}^{B1} \times f_{i,t}^{B1} + w_{i,t}^{B2} \times f_{i,t}^{B2} + ... + w_{i,t}^{Bn} \times f_{i,t}^{Bn}}_{\text{binding metrics}} \quad + \quad \underbrace{w_{i,t}^{D} \times f_{i,t}^{D}}_{\text{discretionary metrics}} \tag{6.2}$$

In Equation 6.2, the firm has committed to individual weights $w_{i,t}^{B1}$, $w_{i,t}^{B2}$, ..., $w_{i,t}^{Bn}$ for n performance metrics at the beginning of the fiscal year. For these n metrics, the executive already knows ex ante with what weight each of them will enter into the calculation of overall target fulfillment $f_{i,t}$. For example, fulfillment rate $f_{i,t}^{B1}$ of metric $B1$ will enter with weight $w_{i,t}^{B1}$. As the firm commits to these weights ex ante, we call them the *binding metrics*. In practice, these binding metrics are often measured with respect to *hard* (quantitative) key performance indicators (KPIs) that can be easily verified. In our analyses, it will be convenient to distinguish between binding *ESG* and binding *non-ESG* performance metrics. For this purpose, we rewrite Equation 6.2:

$$f_{i,t} = \underbrace{w_{i,t}^{B,ESG} \times f_{i,t}^{B,ESG}}_{\text{binding ESG metrics}} \quad + \quad \underbrace{w_{i,t}^{B,nESG} \times f_{i,t}^{B,nESG}}_{\text{binding non-ESG metrics}} \quad + \quad \underbrace{w_{i,t}^{D} \times f_{i,t}^{D}}_{\text{discretionary metrics}} \tag{6.3}$$

where $w_{i,t}^{B,ESG}$ and $f_{i,t}^{B,ESG}$ are the total weight and the joint achievement rate of all binding *ESG* metrics together, and $w_{i,t}^{B,nESG}$ and $f_{i,t}^{B,nESG}$ are defined correspondingly for binding *non-ESG* metrics. In other words, Equation 6.3 shows the relative importance of (binding) ESG and non-ESG performance metrics in the calculation of total target fulfillment $f_{i,t}$.

Besides binding metrics, many firms also include *discretionary* (often called *soft*) metrics. The defining characteristic of these discretionary metrics is that their individual weights are not known to the executive at the beginning of the fiscal year (Maas, 2018).[18] An example

[16] In some firms, directors reserve the right to make adjustments to the formula-based STI payout ex post. We discuss these ex-post adjustments further below.

[17] Payout of realized STI is typically made in cash at or shortly after the end of the performance period.

[18] Ikram et al. (2023) differentiate between objective or formulaic contracts and subjective contracts. They

would be an STI plan that considers the two metrics "Continue to drive strategic initiatives" and "Improve employee satisfaction", but without specifying their relative importance (their individual weights). At the beginning of the year, the executive only knows the joint weight of both discretionary metrics *together*, i.e., $w_{i,t}^{D} = 1 - \sum_{j}^{n} w_{i,t}^{Bj}$. In Equations (6.2) and (6.3), the joint contribution of all discretionary metrics together to overall target fulfillment $f_{i,t}$ enters as the product between their joint weight $w_{i,t}^{D}$ and fulfillment rate $f_{i,t}^{D}$. Economically, these discretionary metrics expose the executive to additional risk as the (supervisory) board / the compensation committee will pick their relative weights only ex post. In other words, the executive cannot be certain how much strong performance on one of these discretionary metrics will actually be rewarded ex post.[19]

Some bonus plans are not additive but connect achievement for different performance metrics multiplicatively. Combinations of the additive and the multiplicative model are also possible – for example, when a subset of metrics is combined additively as in Equation 6.2 before being scaled by a multiplier that measures an additional performance metric. An example would be an STI plan, where the attainment of additively connected financial targets is ultimately multiplied by an individual (non-financial) performance factor. Oftentimes, only ranges of this multipliers or scaling factors are communicated ex ante, whereas the exact values of the multiplier are chosen by the supervisory board only at year-end. Hence, these multipliers can be understood as discretionary metrics.

Finally, realized STI can be subject to deferral conditions, under which STI elements are paid with a certain delay.[20] Usual forms of deferrals include cash deferrals and equity deferrals. In case of cash deferrals, a cash payment is delayed for a certain period of time (e.g., three years) and might also be subject to predefined internal performance thresholds. Equity deferrals represent a deferred payment in the form of real or virtual shares, with the resulting number of shares being blocked for a certain period of time, commonly three years. After expiry of the blocked period, these shares are placed at the beneficiary's disposal.

6.4.2 Long-Term Incentive Pay (LTI)

The lion's share of multi-year incentive plans is represented by the classical long-term incentive plans (LTI) with forward orientation, i.e., variable compensation elements which are based on a performance/vesting period of multiple years. Typical examples include performance cash, performance shares, restricted stock, and traditional stock option plans.

Under performance cash plans, executives are granted a conditional right to a certain cash compensation after the expiry of a predefined performance period, e.g., three to four years. The final amount is determined at the end of the performance period, depending on the achievement of pre-defined performance targets during the performance period. Hence, this constitutes a time- and performance-vesting, long-term bonus plan.

define contracts as objective if a proxy statement clearly specifies the weights (dollar amount of compensation or the percentage) attached to CSR-related metrics.

[19] In our example of the two discretionary metrics "Continue to drive strategic initiatives" and "Improve employee satisfaction", the firm has full discretion to choose the relative importance of both metrics at year-end. For example, the firm is free to concentrate on the first metric at the expense of employee satisfaction.

[20] If part of the payout is deferred, the payout amount less the deferred amount is reported as annual realized STI pay in our data.

Performance shares are time- and performance-vesting stock awards, under which a certain number of shares are conditionally granted to executives at the beginning of a pre-defined performance period. Depending on the achievement of pre-defined performance targets during the performance period, the final number of shares is awarded (which can then increase or decrease). The final amount is delivered either in the form of shares or as an equivalent payment in cash.

Restricted stock plans grant the executive a certain number of real or virtual shares that are blocked for a certain period of time, e.g., four years, and afterward placed at her disposal. Alternatively, at the end of the period, the number of shares may be multiplied by the current share price and the payment is made in cash.

Classical stock options plans grant executives the right (but not the obligation) to purchase shares at a pre-defined strike price. After expiry of the blocking period, the options can be exercised within a pre-specified exercise period, provided that the current share price exceeds the strike price set at the grant date. Payment is then made either in the form of shares or in cash.

6.5 Data

6.5.1 Data Sources

Our sample selection is based on Europe's two leading stock market indices in terms of free-float market capitalization, the EURO STOXX 50 and the STOXX Europe 50. The EURO STOXX 50 tracks the 50 largest and most traded listed companies in the eurozone. The STOXX Europe 50 covers the largest European firms in terms of market capitalization, i.e., it also includes European companies outside the eurozone.[21] Our main sample includes all firms that have been listed for at least ten days on either the EUROSTOXX 50 or the STOXX Europe 50 (or both), between December 31, 2014 and December 31, 2020.[22] We collect data for all firm-years between 2013 and 2020. In case firms enter or exit one of the stock indices within our considered time period 2013 to 2020, we also collect data on the years prior to entry and the years after the exit. If a firm's fiscal year differs from the calendar year, our reporting year refers to the year the company's fiscal year ends. Table E.1 (p. 320) provides an overview about the composition of the indices and the studied firms.

Our empirical analysis builds on multiple data sources. First, information on realized (i.e., ex-post) short- and long-term compensation elements of top executives is provided by the international consulting and board advisory firm Mercer hkp/// group (part of Mercer/Marsh McLennan), a market leader for executive remuneration. This data set comprises information on fixed salaries, annual short-term incentives (STI), multi-year performance elements including deferred compensation, and long-term incentives (LTI), which are typically based on equity. This data is of high accuracy, as the consulting firm offers their executive pay data for benchmarking

[21] As of July 2024, the market capitalization of the EURO STOXX 50 amounts to 4,388 billion Euros and that of the STOXX Europe 50 amounts to 6,266 billion Euros.

[22] A few firms were in of one of the indices for less than ten days, e.g., due to corporate restructuring or (de)mergers. As this is usually only a transitional period, we did not include these companies in our sample. This concerns South32 (8 days, 2015), Uniper (1 day, 2016), Alcon, M&G (both 1 day, 2019), and Siemens Energy (1 day, 2020).

analyses via Europe's leading portal for remuneration data on executive and management board members. Importantly, the data can be compared across countries and companies, due to the use of the consultancies' International Compensation Disclosure Standard (ICDS) methodology.[23]

We complement this data with hand-collected information from companies' annual reports, corporate governance reports, and compensation reports about the various elements of executives' compensation plans. Importantly, we do not limit our data collection to realized compensation at year-end, but also collect information about ex ante contract details at the beginning of the fiscal year. As regards the ex ante contract, we record the target bonus (target STI amount) $TSTI_{i,t}$ and the weights that firms assign to individual metrics (except for discretionary metrics).[24] We further record information on bonus caps and hurdles and whether directors / compensation committees reserve the right to scale STI at their discretion (e.g., via ex post multipliers). As regards the ex post realization at year-end, we collect the realized bonus (realized STI amount) $RSTI_{i,t}$, overall target achievement $f_{i,t}$, specific achievement rates for individual metrics (whenever available), as well as information on whether directors / compensation committees exercised any right to scale STI with an ex post multiplier. Whenever possible, we collect the same information also for executives' LTI plans.[25] As LTI plans are granted over multiple years, information on the number and weights of metrics in LTI plans always refer to the first year of granting such an LTI plan. Note that additional compensation elements such as pension schemes and fringe benefits (company cars, insurances) are not included in our analyses. All compensation data is converted into Euro using year-specific conversion rates provided by the European Central Bank.

We carefully distinguish between ESG and non-ESG performance metrics (as well as between financial and non-financial metrics). Our classification of ESG metrics is based on the ESG framework provided by LSEG Data & Analytics (formerly Thomson Reuters Refinitiv), "one of the largest ESG content collection operations in the world" (LSEG, 2024). This data has been used in more than 1,200 scientific articles, mainly in finance journals (Berg et al., 2020). According to this taxonomy, the "E", "S", and "G" categories are divided into the following 10 subcategories: Emissions, Innovation, Resource use (E), Community, Product responsibility, Human rights, Workforce (S), and Corporate Social Responsibility strategy, Management, Shareholders (G). These classifications can then be subdivided into 375 more detailed metrics. Table E.2 (p. 322) provides an overview of exemplary metrics per category. We use this detailed taxonomy to classify non-financial metrics in executive compensation contracts into non-financial-ESG-related and non-financial-non-ESG-related metrics and to group them accordingly.

We merge our compensation data set with (non-) financial company-level information from LSEG Data & Analytics. This data comprises a series of balance sheet and income statement

[23]See www.boardpay.com and https://boardpay.com/downloads/boardpay_Methodology_ICDS.pdf for details. For all firms that have entered or exited one of the stock indices within our considered time period 2013 to 2020, we hand-collect information on realized compensation elements on the years prior to entry and the years after the exit following Mercer hkp/// group's methodology.

[24]Section E.4 highlights examples of different levels of disclosure of ESG-related metrics in executive STI compensation contracts from company reports.

[25]Like Cohen, Kadach, Ormazabal, and Reichelstein (2023), we observe that reporting granularity varies across firms, countries, regulatory regimes, and over time. Similarly, Bebchuk and Tallarita (2022) document that out of those 53 S&P 100 firms who adopt ESG contracting in 2021, only 27% fully disclose the exact weights of the associated metrics.

items as well as stock price information for the majority of firms in our sample. Corporate governance information comprises data on the fraction of independent and female directors. Corporate carbon emissions are proxied by Scope 1 CO2 and CO2 equivalents emissions. Lastly, we retrieve data on the number of institutional investors and blockholders per firm-year. We define blockholders as large shareholders owning more than 10% of outstanding equity. These firm-level variables are very similar to the ones used in, for instance, Cohen, Kadach, Ormazabal, and Reichelstein (2023) and Hazarika et al. (2022). We winsorize all balance sheet items, realized compensation data (STI and LTI), target bonuses, and executive characteristics (age, tenure) at the 1%-level in each tail of the distribution.

For a detailed description of variable definitions, calculations, and data sources of all main variables, see Section E.1 (p. 316).

6.5.2 Sample Construction

We apply three filters to the raw data ($n = 3,875$ executive-year observations). First, we discard 762 executive-year observations for executives with an annual work period of less than 360 days. In case of executive turnover in the mid of a fiscal year, resulting in only a few months of employment, pay information is hardly comparable to those of incumbent (i.e., non-moving) executives due to special conditions on annual bonus payments in case of leaving and the existence of lump-sum payments such as severance pay. Second, some firms have entered the stock indices as a result of merger and acquisition activities. In these cases, we drop data on the two predecessor companies that have not been part of the stock indices. This results in a drop of 76 executive-year observations. Third, in order to maintain a strongly balanced panel of firms covering all fiscal years from 2013 to 2020, we discard 360 executive-year observations from firms with less than eight firm-years (equivalent to 22 firms).

Our final sample includes 2,677 executive-year observations (584 firm-year observations), derived from 674 distinct executives employed by 73 distinct firms across our eight-year time period. The number of executives within a firm ranges from 1 to 17 with a mean of 4 and a median of 3.

6.5.3 Descriptive Statistics

In Table E.3 (p. 337), we report the composition of our firm sample by industry and headquarter country. Most of our firm-year observations come from the financial services sector, followed by consumers and industrials & materials (Panel A). Unsurprisingly, most firms are headquartered in Germany, the UK, and France (Panel B). We report firm characteristics in Table E.4 (p. 338). Our average sample firm employs 112,332 employees and four executives, has a market capitalization of 59.8 billion Euros, 321.6 billion Euros in total assets, and a ROA (ROE) of 5.4% (15.1%). Female board membership amounts to 31.2% and institutional ownership to 50.3%.

Table E.5 (p. 339), Panel A, reports the main characteristics at the executive level. The average executive is 54 years old and predominantly male, with average tenure in the firm of 6 years. Panel B shows that CEOs account for 576 executive-year observations, followed by CFOs with 399 observations. CHROs and COOs are also relatively frequent (148 and 141 observations each). All other expert C-suite positions (e.g., chief marketing officer, chief sales officer, chief legal officer, etc.) are grouped as *other specialists* (416 observations). Executives heading the management of individual divisions of firms, i.e., executives responsible for certain product lines or geographical markets, account for 825 observations.

Table E.6 (p. 340) provides a first overview of the main characteristics of STI bonus plans in our sample of the largest, listed European firms. The average target bonus (target STI) $TSTI$ equals approximately one million Euros, or 100.6% of annual base salary. STI plans contain an average of 4.6 binding performance metrics, that is, metrics for which firms commit to a specific weight known to the executive ex ante (see Section 6.4, p. 111). Of these binding metrics, 4.2 are non-ESG metrics and only 0.4 are binding ESG metrics. On average, all binding metrics combined have a (total) weight of 75.8% in the payoff function. In the average STI plan, binding ESG metrics have a combined weight of around 2.6%. Discretionary metrics, i.e., metrics for which the firm chooses the weight ex post (at year-end), account for the remaining 24.2% of the payoff function, and an average STI plan is composed of 1.4 discretionary ESG metrics and 3.4 discretionary non-ESG metrics. Furthermore, caps and hurdles in the payoff function are relatively common, with STI bonuses capped, on average, at 177% of base salary. Boards reserve a right to adjust the STI payout, as calculated according to the formula in the STI plan, ex post in more than 70% of the cases, in one-fifth of the cases through ex-post multipliers. Table E.7 (p. 341) shows the corresponding characteristics of LTI plans in our sample.[26] They are characterized by a smaller number of metrics in their year of granting (2.8 binding and 0.8 discretionary metrics in LTI plans compared to 4.6 and 4.8 binding and discretionary metrics in STI plans). Discretionary metrics are less important in LTI than in STI plans, as they have a combined weight of only 8.4% in LTI compared to 24.2% in STI contracts. Focusing on binding ESG metrics, we observe a similar weight of 2.5% in LTI plans (2.6% in STI).

In Table E.8 (p. 342), we break down the different ESG metrics into their environmental, social, and governance components. Both for STI and LTI plans, firms most often include social metrics, followed by governance metrics, whereas environmental metrics are less frequent. In LTI plans, non-ESG metrics are mostly financial performance measures, whereas STI plans have a similar (average) number of financial and non-financial non-ESG metrics.

Finally, Table E.9 (p. 343) shows statistics for STI fulfillment (achievement) rates as well as realized compensation at year-end. On average, overall target achievement amounts to 104%, suggesting that executives hit close to the target set at the beginning of the fiscal year. However, the standard deviation is large at 40%. The average base salary is 1,093,465 Euros and average realized bonus (realized STI) $RSTI$ amounts to 821,716 Euros, i.e., 84% of base salary. The average realized LTI amount $RLTI$ is 1,726,691 Euros. Realized STI is partially deferred for about half our sample. In case of deferral, 53% of the STI amount is deferred over an average deferral period of 3 years. Although boards reserve the right for discretionary adjustments in more than two-thirds of cases, they exercise these discretion rights in only 13% of all cases.

[26] Note that we do not report target values for LTI plans for several reasons. First, LTI target values are reported in a minority of cases, with missing information being present particularly in earlier years. Additionally, LTI details are sometimes not specified at the individual level, but only for the entire executive team. Lastly, it often occurs that multiple LTI plans in the same firm overlap over time, as each of these plans have different vesting periods, which makes it difficult to determine LTI target achievement across time.

6.6 Empirical Analysis

Our empirical analysis consists of three parts. In Section 6.6.1, we study the relevance of ESG metrics in executive STI contracts, analyzing the reported numbers and weights of different ESG metrics as well as their contribution to executives' pay risk. In Section 6.6.2, we show how the design of ESG metrics and their contribution to pay risk vary across different industries and firms and in Section 6.6.3, we do a similar comparison across different executive positions. We focus on STI contracts because of better data availability (see footnote 26) and because previous research has shown that ESG metrics are more often used in STI vs. LTI schemes (Cohen, Kadach, Ormazabal, & Reichelstein, 2023; Homroy et al., 2023; Walker, 2022; Willis Towers Watson, 2023).

6.6.1 Relevance of ESG Pay

In this section, we show that many firms report a large number of ESG metrics that lack a material weight in the calculation of total STI (Section 6.6.1.1), that variation in ESG target achievement is low, and that ESG metrics account for only a small fraction of total STI variance (Section 6.6.1.2). We compare these initial findings against Hypothesis 1 in Section 6.6.1.3.

6.6.1.1 Prevalence of ESG Performance Metrics in Listed European Firms

To measure the adoption of ESG pay, most papers check whether a given compensation plan is tied to at least one ESG criterion (e.g., Carter et al., 2023; Cohen, Kadach, Ormazabal, & Reichelstein, 2023; Hazarika et al., 2022; Ikram et al., 2023). We apply this "extensive margin approach" to 674 executives of 73 firms that are listed on the EURO STOXX 50 or the STOXX Europe 50 between 2013 and 2020. Panel A of Figure E.6 (p. 330) shows that in 2013, only 40% of executives have STI pay linked to ESG metrics, which increases to approximately 60% in the second half of our sample period (dashed black line). By 2020, nearly 60% of executives have at least one social metric, just over 30% have one or more governance metrics, and only 19% have environmental criteria. In contrast, all executives have one or more non-ESG performance metrics across the entire sample period (solid line at 100%).

We further break down each of the three main ESG categories into their subcategories (see Table E.2, p. 322, for definitions and examples of these subcategories). Panel B of Figure E.6 (p. 330) shows that emissions reduction metrics are the most common environmental criteria, increasing from 4% to 15% of executives by 2020, with the most significant rise from 2018 to 2020, i.e., precisely at a time when the price of EU carbon permits quadruples (see Figure E.13, p. 350). Metrics for resource use, such as water and energy efficiency, increase from 0% in 2013 to an executive share of 7.5% in 2020. Innovation and green R&D metrics remain uncommon and are found for only 2.4% of executives by 2020.

For social metrics, Panel C reveals that workforce-related and product quality metrics are the most common, reaching executive shares of about 45% and 30%, respectively, by 2020. Social criteria that consider possible externalities on local communities are much rarer, with a share of about 10%, and not a single executive has any metrics explicitly related to human rights. Panel D shows that governance metrics related to CSR strategy and management are the most prevalent, covering about 20% of executives in 2020, where the strong increase over our sample period is likely driven by an anticipation of new regulations towards improved corporate governance, such as board independence and board diversity.

Overall, Figure E.6 (p. 330) suggests that while a substantial share of executives have at least one ESG performance metric, firms tend to prefer criteria that may improve financial performance, such as emissions and workforce-related criteria, over metrics addressing broader externalities, such as human rights.

While the extensive margin approach in Figure E.6 (p. 330) tells us about the *presence* of ESG in STI, it remains silent about its *importance* relative to non-ESG criteria. In a first attempt to overcome this limitation, Figure E.7 (p. 331) plots the *number* of ESG versus non-ESG metrics per executive. Panel A shows that, on average, executives have four times more non-ESG than ESG metrics, and Panel B reveals that the average executive has three times more social metrics than governance-linked metrics and very few environmental criteria.[27] Together, Panels A and B suggest that ESG metrics, particularly social outcomes, carry only about one-fourth the importance of non-ESG criteria. Although this is a step forward compared to Figure E.6 (p. 330), there remain two limitations to this analysis.

First, it remains unclear if ESG metrics are binding and how heavily they factor into year-end performance assessments. As explained in Section 6.4 (p. 111), *binding* metrics have a pre-set weight known to the executive, whereas *discretionary* metrics lack an explicit ex-ante weight and may even be disregarded by the firm ex post. Interestingly, Panel C of Figure E.7 (p. 331) shows that especially the number of discretionary ESG metrics increases during our sample period, whereas binding ESG metrics only increase slightly after 2018. A second limitation is that firms can inflate the perceived importance of ESG-linked pay by splitting a given ESG outcome into several criteria, even if they lack substantial weight or are non-binding. To overcome these limitations, Figure E.8 (p. 332) plots the joint weight that firms give to binding ESG metrics together. We find that the total weight of all binding ESG metrics combined is, on average, less than 2% in 2013. Even after a strong increase between 2018 and 2020, this ESG weight reaches only 5% by 2020.

Overall, the extensive margin analysis reveals little about the actual importance of ESG pay. By 2020, although 60% of executives had at least one ESG metric, ESG metrics were still outnumbered by non-ESG metrics, were often discretionary, and, when binding, carried modest weights in STI calculations.

6.6.1.2 Contribution of ESG Metrics to Pay Risk

Whether an executive has incentives to work towards a given ESG performance metric depends on two factors. First, the metric must carry substantial weight in the performance assessment; otherwise, the executive is likely to disregard it unless intrinsically motivated (in which case incentive pay is unnecessary). Second, achieving the target should be contingent on the executive's

[27] We plot average numbers of ESG metrics per executive because plotting median numbers would not reveal any new information. Figure E.6 (p. 330) already shows that the median executive has zero metrics in most ESG categories.

actions. As discussed in Section 6.3 (p. 108), some firms might want to greenwash executive pay and design ESG metrics in a way that makes their achievement automatic. In this case, variation in the achievement rates of ESG metrics would be zero, making ESG-linked STI effectively part of fixed salary with no incentive power.[28] Thus, greenwashing implies that ESG criteria subject executives to minimal pay risk, due to small, immaterial weights of ESG metrics and/or stable ESG achievement rates.

Our goal is to estimate the contribution of ESG performance metrics to the pay risk posed by the STI contract of executives. To this end, we decompose total STI variance into its different components. First, we substitute Equation 6.3 into Equation 6.1 to rewrite total realized STI:

$$\begin{aligned} RSTI_{i,t} &= TSTI_{i,t} \times w_{i,t}^{B,ESG} \times f_{i,t}^{B,ESG} + TSTI_{i,t} \times w_{i,t}^{B,nESG} \times f_{i,t}^{B,nESG} + TSTI_{i,t} \times w_{i,t}^{D} \times f_{i,t}^{D} \\ &= RSTI_{i,t}^{B,ESG} + RSTI_{i,t}^{B,nESG} + RSTI_{i,t}^{D}, \end{aligned} \tag{6.4}$$

where $RSTI_{i,t}^{B,ESG}$ and $RSTI_{i,t}^{B,nESG}$ are the parts of total realized STI that the executive receives for the achievement of binding ESG and non-ESG metrics. As explained in Section 6.4 (p. 111), discretionary ESG and non-ESG metrics lack individual weights and are only assessed jointly. Hence, it is not possible to split $RSTI_{i,t}^{D}$ into an ESG and a non-ESG part. Finally, we can decompose the variance of total realized STI into the following components:

$$\begin{aligned} Var\left(RSTI_{i,t}\right) = {} & Var\left(RSTI_{i,t}^{B,ESG}\right) + Var\left(RSTI_{i,t}^{B,nESG}\right) + Var\left(RSTI_{i,t}^{D}\right) \\ & + 2Cov\left(RSTI_{i,t}^{B,ESG}, RSTI_{i,t}^{B,nESG}\right) + 2Cov\left(RSTI_{i,t}^{B,ESG}, RSTI_{i,t}^{D}\right) \\ & + 2Cov\left(RSTI_{i,t}^{B,nESG}, RSTI_{i,t}^{D}\right) \end{aligned} \tag{6.5}$$

Table E.10 (p. 344) shows this variance decomposition for different samples. We report total STI variance $Var\left(RSTI_{i,t}\right)$ in column 2. In Panel A, this total variance includes both within-executive and between-executive STI variation. In the full sample of 1,076 executive-year observations with information on all necessary variables, binding ESG metrics account for only 1.9% of this total STI variance (column 4). Interestingly, this variance share is smaller than the weight of 3.5% that these metrics take in the overall performance assessment at year-end (column 3).[29] Hence, binding ESG metrics contribute less to total STI variance than their weight may suggest.

Binding non-ESG performance metrics have by far the largest variance share at 88.1% (column 6), which slightly exceeds their weight of 87.5% (column 5). Hence, binding non-ESG metrics contribute 46 times (88.1/1.9) more to overall STI variance than binding ESG metrics. The contribution of the covariance between binding ESG and non-ESG metrics is positive with a variance share of 4.4% (column 9), suggesting that both are positively correlated. Finally, all discretionary metrics together account for another 9.3% of total STI variance (column 8).

[28] For example, greenwashing firms can set the ESG metric so low that the executive is sure to reach the highest possible level of target achievement. Alternatively, firms might be able to hide the executive's true performance from the public (e.g., if the performance metric is intransparent and hard to verify) and then always declare full target achievement. In both examples, the executive is guaranteed to get the highest possible bonus for ESG performance.

[29] The weight of 3.5% is calculated as the sample average among the 1,076 executive-year observations.

A potential issue with our analysis could be that some firms (i.e., their supervisory boards and/or compensation committees) do not calculate realized STI based on weights and achievement rates alone but reserve the right to make a final adjustment to total STI at their discretion (see Section 6.4, p. 111). A priori, it is unclear how such an adjustment to total STI should be attributed to the different performance metrics. However, Table E.10 (p. 344) shows that the variance shares change only marginally if we drop these firms with supervisory board discretion and focus on the remaining subset of 926 observations.

As shown in Figure E.6 (p. 330), 40% to 60% of the executives in our sample (black, dashed line in Panel A) do not have a single ESG criterion in their STI. For these executives, the contribution of ESG metrics to total STI variance is (mechanically) zero. After we drop the corresponding observations and retain only executives with at least one ESG metric, the weight of binding ESG metrics increases from, on average, 3.5% to 8.5% (column 3). Strikingly, the variance share of these metrics increases much less and remains low at 3.5% (column 4), indicating again that binding ESG metrics contribute less to total STI variance than their explicit weight suggests, even in a sample of executives whose contracts are all related to at least one ESG metric.

As explained above, Panel A of Table E.10 (p. 344) considers both within-executive and between-executive STI variance. However, the pay risk that STI poses to an individual executive is better measured by within-variation alone, that is, by how much the STI of a given executive varies over the years she works at a given firm. To eliminate all between-executive variation in STI, we regress $RSTI_{i,t}^{B,ESG}$, $RSTI_{i,t}^{B,nESG}$, and $RSTI_{i,t}^{D}$ in Equation 6.4 on interacted executive $\times$ firm fixed effects $\theta_{i,f}$.[30] Then we redo the variance decomposition for the regression residuals, which only exhibit within-executive STI variation.

Panel B of Table E.10 (p. 344) reports the results of this new decomposition. In the largest available sample with 1,074 observations, binding ESG metrics account for only 1% of within-executive STI variance (column 4). This implies that the variance share of binding ESG metrics is three times smaller than the weight of 3.5% that these metrics take in the assessment of overall performance at year-end (column 3). Even if we focus only on executives with at least one ESG metric, this variance share remains small at 2.5%, and binding ESG metrics explain 26 times (64.9/2.5) less of within-executive STI variance than binding non-ESG metrics.

Overall, our variance decomposition in Table E.10 allows us to draw two conclusions. First, it shows that ESG performance metrics contribute very little to the pay risk that STI poses to executives. Focusing on within-executive STI variance (Panel B), we estimate that depending on whether we consider only binding or also discretionary metrics, the risk contribution of ESG is somewhere between 1% (column 4 of Panel B) and 8.7% (column 4 + column 8) in the full sample, and between 2.5% and 17.7% in the subsample of executives with at least one ESG metric.

Second, the variance decomposition also gives us an idea of *why* ESG contributes so little to the overall pay risk posed by executives' STI bonus plans. Of course, the main reason is the small weight of ESG metrics in the year-end assessment of overall performance. Beyond this, it also seems likely that ESG criteria have very stable target achievement rates, since their contribution

[30] We use the interaction to account for cases in which the same person subsequently works for different firms.

to overall STI variance falls well short of their already small weights (compare columns 3 and 4 of Panel B). Indeed, additional analyses show that the year-end assessment of overall performance exhibits less annual variation if ESG criteria take a larger place in the executive's STI contract. In Table E.11 (p. 345), we calculate the annual variation (S.D.) of the overall achievement rate $f_{i,t}$ for each of the 65 CEOs in our sample and regress it on numbers (columns 1-3) and weights of ESG and non-ESG metrics (columns 4-6). Regardless of the set of controls, overall performance achievement is less volatile for executives with more *ESG* performance metrics and with higher weights on these metrics, whereas *non-ESG* metrics are not associated with higher performance volatility. Hence, ESG criteria seem to pose relatively little pay risk, both due to immaterial metric weights and also because they seem to reduce variation in overall performance achievement.

6.6.1.3 Interim Conclusion for Hypothesis 1

Part (a) of Hypothesis 1 predicts that under incentive contracting, ESG metrics have a *material* weight in the calculation of executives' performance pay, and that ESG target achievement is *not* automatic. In contrast to this prediction, Sections 6.6.1.1 (p. 118) and 6.6.1.2 (p. 119) show that most ESG metrics are either discretionary (firms can ignore them at year-end) or that firms commit only to very small weights ex ante. As a consequence, they explain only an immaterial portion of total executive pay risk posed by STI plans. Overall, the fact that firms report an increasingly large number of ESG metrics (for up to 60% of their executives), albeit with questionable incentive power, seems more consistent with greenwashing (Part (b) of Hypothesis 1). Similarly, the small ESG weights also seem inconsistent with rent extraction (Hypothesis 4).

6.6.2 Industry and Firm Characteristics

In this section, we study heterogeneity in the prevalence and design of ESG-linked STI plans between different industries (Section 6.6.2.1, p. 122) and firms (Section 6.6.2.2, p. 123). We observe that historically more polluting and energy-reliant industries, as well as firms with noisier stock prices, mostly use *binding* ESG performance metrics. By contrast, discretionary (non-binding) ESG performance metrics are especially common in the financial sector and in large firms with institutional but dispersed ownership and with more independent directors. We compare these findings to Hypothesis 2 in Section 6.6.2.3 (p. 125).

6.6.2.1 Industry Differences

In Figure E.9 (p. 333), we begin by plotting the numbers of discretionary and binding ESG performance metrics per executive by industry. We observe a striking difference in the prevalence of both types of performance metrics, that is, binding and discretionary ESG metrics are common in different industries. In particular, *discretionary* metrics are most popular in the financial sector, and least common in the energy & utilities sector, whereas the opposite is true for *binding* ESG metrics. Financial firms rapidly increase the number of discretionary ESG metrics from 1 to approximately 4 between 2013 and 2020. These same financial firms have no binding

ESG metrics until 2018, and only about 0.75 binding metrics per executive in 2020. By contrast, in the energy & utilities sector, companies start with only 0.5 discretionary ESG metrics in 2013 and further reduce this number to about 0.25 metrics in 2020. For binding ESG metrics, energy & utilities take the opposite trend and increase the number of binding metrics per executive from about 0.25 to 1.75 in 2020.

Unsurprisingly, the industries that favor binding ESG performance metrics over discretionary ones also have the highest weights for these binding metrics in the year-end calculation of total STI. Figure E.10 (p. 334) shows that industrial & materials, ICT, and energy & utility firms commit to weights that are between 2 and 3 times larger (depending on the year) than in the financial, consumer, and health care industries. These higher ESG weights also translate into a larger contribution of binding ESG metrics to the overall pay risk that STI poses to executives.

Panel A of Table E.12 (p. 346) shows that binding ESG metrics contribute 6.1% to total STI variance in energy & utilities (column 4), but only 0.9% in finance. When considering only within-executive STI variance in Panel B, those variance shares decrease, but the difference between both sectors remains qualitatively unchanged. Interestingly, the covariance terms between ESG and non-ESG metrics in column 9 are also very different in both sectors. In finance, the correlation between ESG and non-ESG performance pay is close to zero – the covariance term is only 6.8% in Panel A and 3.1% in Panel B. By contrast, this covariance term between ESG and non-ESG is larger in energy & utilities firms than in any other industry – i.e., 27.1% in Panel A and 14.2% in Panel B. One possible explanation could be that ESG and non-ESG performance have an especially strong, positive relation in the energy & utilities sector (doing well by doing good).

6.6.2.2 Firm Characteristics

Next, we move from a broad industry comparison to fixed effects regressions that allow us to study firm heterogeneity *within* industries. To this end, we estimate the following model:

$$y_{i,t} = \alpha + \beta_1 X^{firm}_{f,t} + \beta_2 X^{exec}_{i,t} + \mathit{IndustryFE} + \mathit{ExecPositionFE} + \mathit{YearFE} + \epsilon_{i,t} \tag{6.6}$$

where we use different dependent variables $y_{i,t}$ to measure the adoption of ESG performance metrics in the STI compensation plan of executive i in year t. In this section, we focus on the vector of firm characteristics $X^{firm}_{f,t}$ to gauge possible differences between firms conditional on industry fixed effects. As firm characteristics, we include a list of variables similar to that in Cohen, Kadach, Ormazabal, and Reichelstein (2023), subject to a few conscious changes. In particular, to measure a firm's scope-1 emissions of CO_2 equivalents, we use the *historical* (log) average in the five years *before* the start of our sample period (hence, we use the years 2008-2012), and we use the same pre-sample years to measure firms' stock-to-accounting volatility (see Section E.1, p. 316). In both cases, the historical values should address potential concerns about reverse causality. Besides industry fixed effects, Equation 6.6 further includes fixed effects for time and for different executive positions (CEO, CFO, etc.), as well as a vector of different executive characteristics $X_{i,t}^{exec}$, which we will analyze in Section 6.6.3 (p. 125). Standard errors are clustered at the firm level.[31]

Table E.13 (p. 347) shows regressions for different dependent variables $y_{i,t}$. We study the use of discretionary ESG performance metrics in executives' STI contracts in columns 1 to 4 and the use of binding ESG performance metrics in columns 5 to 8. Previous research mostly uses a binary variable that equals one if a compensation plan includes at least one ESG performance metric, and zero otherwise. We adopt this approach in columns 1 and 5. Then we refine the analysis and count the (log) number of ESG metrics in columns 2 and 6. To gauge the relative importance of ESG and non-ESG performance metrics, we standardize the number of ESG metrics by the total number of reported (ESG and non-ESG) metrics in columns 3 and 7. Finally, we use the weight of binding ESG metrics in column 8 and the weight of all (ESG and non-ESG) discretionary metrics together in column 4.[32]

Like in our industry analysis, we observe again a striking difference between firms that use discretionary and firms that use binding ESG performance metrics. In Table E.13 (p. 347), the only statistically significant firm characteristic that has the same (positive) coefficient sign for both types of ESG metrics is the binary variable *Emissions policy*, which identifies firms that have officially enacted some policy to reduce emissions. For all other firm characteristics, the coefficient estimates are either statistically insignificant or have opposing signs for discretionary and binding ESG metrics, which suggests again that both types of metrics serve very different purposes.

Our first, maybe surprising, result in Table E.13 is that *discretionary* ESG performance metrics are significantly *less* common in firms with historically higher CO_2 emissions (columns 1 to 4). This contrasts with Cohen, Kadach, Ormazabal, and Reichelstein (2023), who predict that "more polluting firms have a higher incentive to improve ESG performance (they face a higher cost for their emissions and could suffer from stranded assets) (p.824)." One possible explanation could be that *discretionary* ESG metrics, which firms are free to ignore if they wish to, are simply not suited to provide strong incentives. Indeed, our industry analysis in the previous section has already shown that sectors with a large environmental footprint, like energy & utilities, shun discretionary ESG metrics. Instead, discretionary ESG metrics are more common in larger (high *Log(total assets)*) and more mature (high *Log(book-to-market ratio)*) firms, which tend to be more visible and likely attract more public scrutiny (Cohen, Kadach, Ormazabal, & Reichelstein, 2023). Similarly, discretionary ESG metrics tend to be associated with more independent directors (high *Board independence*) and larger institutional but dispersed equity ownership (high *Institutional ownership* but low *Block ownership*) – that is, with variables that likely capture stronger ESG pressure.[33] Interestingly, we observe a very different relationship with ownership structure for *binding* ESG metrics, as the coefficients of *Institutional* and *Block ownership* flip signs between columns 1 to 4 and columns 5 to 8, suggesting that blockholders

[31] We do not cluster by year because we only have 13 years of data, whereas the conventional threshold for the cluster-robust covariance matrix to converge is 50 clusters.

[32] We cannot use the weight of discretionary ESG metrics alone because as explained above, firms do not commit to explicit weights for individual, discretionary metrics. They only report one overall weight for all discretionary metrics together, i.e., the combined weight of discretionary ESG and non-ESG metrics.

[33] For example, Azar et al. (2021) and Cohen, Kadach, and Ormazabal (2023) show evidence that institutional investors care about ESG performance.

have different preferences over binding vs. discretionary ESG performance metrics than other institutional owners.[34]

Finally, we observe that firms with noisier stock prices, as measured by a high historical stock-to-accounting volatility (Li & Wang, 2016), use *binding* ESG performance metrics more often (columns 5 to 8), consistent with the prediction that firms rely on these metrics when their equity is too noisy as a signal for managerial performance.[35] Strikingly, the coefficient of *Historical stock-to-accounting volatility* is only significant for binding but not for discretionary ESG metrics.

6.6.2.3 Interim Conclusion for Hypothesis 2

Hypothesis 2 suggests that binding ESG metrics with actual incentive power are more prevalent in sectors where ESG performance enhances shareholder value and in firms where equity alone is an unreliable indicator of managerial performance. Consistent with these predictions, binding ESG metrics are more common in firms with a high environmental footprint, i.e., in firms with significant environmental costs, such as carbon pricing, as well as in firms with noisier stock prices. In contrast, discretionary ESG metrics – those without clear incentive power – are common in larger, high-visibility firms, especially in finance, where public attention is high. Overall, these findings suggest that binding and discretionary ESG metrics serve different purposes and are thus prevalent in different firms, depending on whether the focus is on genuine incentive provision or on greenwashing.

6.6.3 Executive Characteristics

In this section, we study heterogeneity in the prevalence and design of ESG-linked STI plans between different executive positions. We find that ESG-linked STI tends to be more prevalent among generalists like CEOs than among specialists. We find no evidence that firms tailor ESG performance metrics to the specific tasks that a given executive must perform in her job, nor that they tailor metrics to demographic characteristics like age or gender (Section 6.6.3.1, p. 125). We compare these findings to Hypothesis 3 in Section 6.6.3.2 (p. 127).

6.6.3.1 Which Executives Have Which Types of ESG Performance Metrics?

Figure E.11 (p. 335) shows the numbers of discretionary and binding ESG performance metrics by executive position. Between 2013 and 2016, metric numbers are roughly the same across all different positions. However, after 2016, we observe an increase in discretionary ESG metrics (Panel A) for CEOs, COOs, and for our catch-all category "Other specialist", which groups all executive positions with very few observations. After 2016, CEOs and COOs also experience a

[34] Cohen, Kadach, Ormazabal, and Reichelstein (2023) use *Controlling shareholder*, which equals one if one shareholder owns more than 50% of the equity. In our European sample, only Volkswagen AG has a controlling shareholder. Therefore, we include *Block ownership*, defined as the fraction of equity held by investors that hold at least 10% of total outstanding shares, instead.

[35] The assumption that variance in stock returns proxies for the signal-to-noise ratio of equity as a measure of executive performance is common in the literature and has been shown to explain variation in CEO incentives across firms (e.g., Aggarwal & Samwick, 1999; Garen, 1994; Garvey & Milbourn, 2003; Himmelberg et al., 1999; Jin, 2002; Lambert & Larcker, 1987; Li & Wang, 2016).

strong increase in the number of binding ESG metrics (Panel B). Similarly, Figure E.12 (p. 336) shows that after 2016, firms commit to increasingly large ESG weights in the STI of CEOs, COOs, and also CFOs. Consistent with both figures, the variance decomposition in Table E.14 (p. 348) confirms that CEOs, COOs, and CFOs have larger variance shares for binding ESG performance metrics than other executives, although these variance shares remain low at an absolute level.[36] Overall, the evidence suggests that ESG-linked pay is more common among generalist executives than among specialists, as CEOs, COOs, and (to a lesser degree) also CFOs tend to have a broader spectrum of tasks than most other executives.[37]

While specialists have fewer ESG metrics than generalists, firms can still tailor these metrics to the specific tasks each specialist is responsible for. However, the specialists' executive positions indicate only the broad corporate policies they oversee, and our data do not clarify how these should correspond to specific ESG performance metrics. Nonetheless, in two cases, this alignment appears relatively straightforward. For example, reducing a firm's environmental footprint often involves developing new technologies and replacing stranded assets. Therefore, we would expect chief technology officers (CTOs) to have more environmental ESG metrics, particularly those linked to emissions. Similarly, a firm aiming to improve working conditions and employee well-being will likely assign this task to its CHRO. As a result, CHROs are expected to have social ESG metrics, especially workforce-related metrics, more often than other executives. To test these predictions, we estimate the following regression:

$$y_{i,t} = \alpha + \beta_1 CTO_{i,t} + \beta_2 CHRO_{i,t} + Firm \times \ Year \ FE + \epsilon_{i,t} \tag{6.7}$$

where the dependent variable $y_{i,t}$ equals one if the STI of executive i is linked to a given ESG metric in year t, and zero otherwise. The regressors $CTO_{i,t}$ and $CHRO_{i,t}$ are also binary and equal one if executive i is a CTO or, respectively, a CHRO, and zero otherwise. We condition the estimation on firm fixed effects interacted with year fixed effects. The regression sample is restricted to CTOs, CHROs and CEOs. Hence, the coefficient β_1 tells us whether a given CTO is more or less likely to have metric $y_{i,t}$ than the CEO of the same firm in the same year. Similarly, β_2 compares the propensity of a given metric between the CHRO and the CEO in the same firm and year. In keeping with our prediction of "tailoring" metrics to positions, we expect β_1 to be positive for environmental metrics, and β_2 to be positive for social metrics.

In Table E.15 (p. 349), we report the coefficient estimates for Equation 6.7. In columns 1 and 2 of Panel A, the dependent variable equals one if executive i has any environmental metric. The coefficients are all negative and significant, indicating that CTOs are less and not more likely than CEOs to have environmental metrics. Moreover, CTOs are not even more likely to have environmental metrics than CHROs. This result remains unchanged when we look specifically at metrics to reduce emissions (columns 3 and 4 of Panel A). Panel B further shows that social

[36] This difference is stronger in Panel A (total STI variance decomposition) than in Panel B (only within-executive variation).

[37] CEOs are responsible for overall firm performance, for integrating the different corporate policies into a unified strategy, and for shaping the firm's strategic vision. COOs act as the CEO's right hand, overseeing all daily, internal operations. CFOs nowadays also require, besides expert knowledge in finance and financial reporting, a high-level understanding of firm-wide strategy and broader management responsibilities, as their financial decisions affect all other corporate policies (Brickley et al., 2017).

metrics in general (columns 1 and 2) and workforce-related metrics in particular (columns 3 and 4) are not more common among CHROs than among CEOs. These findings run counter to our prediction that firms tailor ESG metrics to the tasks of different specialists.

Finally, we return to our initial regression model in Equation 6.6 and study the coefficient estimates for executives' demographic characteristics $X_{i,t}^{exec}$. Table E.13 (p. 347) shows coefficients for executive tenure and age (in years) and a binary variable equal to one if the executive is female. Again, the overwhelming majority of coefficient estimates for these three variables is statistically insignificant across the different dependent variables. Overall, we see no evidence of metric tailoring, neither to executives' positions nor to their individual characteristics.

6.6.3.2 Interim Conclusion for Hypothesis 3

Hypothesis 3 predicts that under incentive contracting, ESG performance metrics are tailored to the job of the executive and are more common among specialists than among generalists. By contrast, if firms include ESG metrics to greenwash STI (and not to incentivize executives), then ESG metrics should be most common among the more visible executives that attract most public scrutiny. The evidence above is more consistent with the greenwashing rational of Hypothesis 3. First, we see no evidence that firms carefully tailor ESG performance metrics to the executives' different tasks. Second, ESG metrics are most common among CEOs, CFOs, and COOs, whose compensation plans likely attract most public attention and scrutiny. If firms want to greenwash executive pay, it would make sense that they start with those high-ranking and more visible generalists.

6.7 Conclusion

In conclusion, this study reveals significant heterogeneity in the role of ESG metrics in executive compensation across industries, firms, and executive positions. While ESG metrics are increasingly prevalent, their integration into executive pay often lacks material weight and incentive power. Discretionary ESG metrics dominate in financial firms and large, visible companies, raising concerns of greenwashing rather than genuine incentive alignment. In contrast, firms in energy-intensive and high-polluting industries, as well as those with highly volatile stock prices, tend to adopt binding ESG metrics with more substantial weights. This variation highlights the complex and evolving role of ESG metrics in executive compensation, with some firms adopting ESG-linked pay as part of a broader incentive framework, while others use it more for external signaling. Standardization in ESG-related compensation practices may increase with regulatory changes and rising transparency, providing a clearer framework for assessing the true impact of ESG criteria on executive behavior.

Chapter 7

Conclusion

Recent technological advances have profoundly reshaped modern life.[1] Job applications are rarely submitted on paper, searching for a roommate no longer involves posting a note on a university bulletin board, romantic relationships are often initiated online, and intimate personal information is now just a few clicks away rather than requiring a private investigator. Consequently, there is a pressing need for rigorous causal evidence regarding the behavioral outcomes of digitization. One major outcome of this shift is the pervasive influence of social media and other digital platforms. Evaluating the consequences, challenges, and behavioral changes driven by formal and informal online settings and *non-traditional* information from social media and other online (matching) platforms is a crucial task for researchers seeking to inform the scientific community, society, and policymakers – particularly with regard to unequal treatment of market participants with different characteristics and backgrounds.

Moreover, growing sustainability concerns have led the public and institutional investors to demand that listed companies incorporate *non-traditional* performance criteria, such as environmental and social metrics, into remuneration packages to promote sustainable businesses and long-term viability. However, a critical question arises: Are these firms incentivized to adopt clearly measurable, adequately weighted, and binding performance metrics that shape executives' behavior and firm strategy, or to use vague, discretionary, and easily attainable goals?

These issues and their practical consequences require thorough scientific investigations to yield effective policy recommendations and keep the public informed about how technological and behavioral shifts shape economic and social outcomes. This thesis seeks to make a significant contribution in this regard by investigating how novel or non-traditional sources of information shape economic and social decision-making.

This thesis contributes to several strands of the literature: By employing large-scale field experiments, Chapters 2–4 provide robust causal evidence on how minority applicants are penalized in everyday social and economic life – and the conditional role of stereotypes and social media information in amplifying or mitigating these biases. Through controlled manipulations of social media profiles, these studies offer novel insights into how digital self-presentation can alter online and offline outcomes, highlighting both the promise and pitfalls of additional online transparency. Similarly, Chapter 5 provides causal evidence on the effects of different personality

[1] Latest advances in AI not even considered.

traits on the same outcomes providing data on real-world behavioral responses as a result of an individuals' social media profile. Chapter 6 adds to the growing literature on ESG by providing granular evidence of whether, how, and why ESG metrics affect executive decision-making. It calls attention to the risk of symbolic adoption in the face of public pressure, reinforcing the need for more robust regulatory or stakeholder oversight.

Discrimination, Information, and Social Media Chapter 2 examines whether social media profiles free of minority stereotypes can mitigate discriminatory behavior by potential roommates in the shared housing market. The study randomly varies applicants' ethnicity, gender, and whether they include a link to a social media profile devoid of any cultural or religious signals, thereby breaking with existing stereotypes for minority applicants and prompting an update of initial (stereotyped) beliefs. Results indicate an ethnic gap in callbacks of 16 percent (significant at the one percent level), i.e., ethnic minority applicants have to send out 16 percent more applications on average to receive an invitation for a viewing compared to majority applicants. While discrimination is sizable in the absence of a social media profile (32 percent), it drops to only 4 percent – almost eliminating the ethnic gap – when a profile free of minority stereotypes is included. Moreover, the minimal remaining gap is smaller for male applicants, a group typically subject to higher levels of discrimination than females – especially in housing markets (Bertrand & Duflo, 2017; Bosch et al., 2010; Carlsson & Eriksson, 2014; Carpusor & Loges, 2006). Overall, these findings indicate the potential of carefully curated social media information to substantially reduce ethnic discrimination.

Chapter 3 examines how minority stereotypes, expressed through "visual narratives" signaling ethnic background, cultural identification, and religious affiliation – shape unequal treatment in personal social network formation and the informal housing market. These visual stereotypes are applied to the fictitious social media profiles introduced in Chapter 2. By randomly varying applicants' ethnicity and whether an application includes a link to a social media profile with or without visual minority stereotypes, this study evaluates the impact of visual stereotypes on ethnic discrimination. The findings reveal that applicants with social media profiles that show minority stereotypes are 69 percent less likely to be accepted as friends and 71 percent less likely to be considered as potential roommates. Visual minority stereotypes reduce acceptance and callback rates for all ethnic groups. However, the negative effect is significantly more pronounced for minority applicants.

Contributing to the growing economic literature of salience (Bordalo et al., 2022; Bordalo et al., 2020), Chapter 4 explores how an application's salience – measured by its position in the roommates' inbox – affects discriminatory behavior. The study randomly varies whether an application is sent from a premium account (requiring a monthly subscription) that places the application near the top of the inbox. The findings show that premium status modestly increases callback rates for minority *and* majority applicants. However, it does not eliminate the persistent ethnic penalty of minority applicants.

Overall, the results from Chapters 2 through 4 suggest that social media profiles free of minority stereotypes effectively reduce bias, that inequalities in informal settings and markets remain largely unaddressed, and that attention-based mechanisms or increased salience do not substantially reduce ethnic discrimination. However, these experiments primarily focus on ethnic cues (names, photos). Real-world social media profiles typically include multiple overlapping signals – such as race, age, and social class – which may amplify the complexity of bias in more diverse contexts. Future research could expand on this limitation by examining the intersection of multiple identity dimensions (e.g., ethnicity, gender, socioeconomic background, sexual orientation), offering a more nuanced understanding of how layered stereotypes impact decision-making.

Further research is also needed into how online platforms might curb discrimination – whether through anonymization, structured profile designs, or curated feedback mechanisms. Identifying platform features that effectively reduce biases at scale could not only inform both corporate strategies and public policy, but also substantially reduce inequality and maybe even positively affect the long-lasting lock-in effect of minority members (see Chapter 1).

Despite these open questions, Chapters 2 to 4 demonstrate the critical need for targeted interventions to address implicit biases in both digital and real-world domains. Addressing subtle, often overlooked forms of discrimination within informal social and economic settings beyond the realm of formal labor or professional markets remain essential to reduce inequality.

Personality and Social Media Chapter 5 shifts from ethnic discrimination to visual personality cues. Through multiple large-scale online pilot experiments, this research develops visual representations of three personality traits which are substantially affecting individual outcomes in organizational and other settings: agreeableness/emotional stability and conscientiousness. We then conduct two field experiments to assess how these personality signals, embedded in fictitious social media profiles, affect friend acceptance rates by social media users and callback rates by potential roommates. The results indicate that profiles signaling high agreeableness/emotional stability receive significantly higher acceptance and callback rates than those signaling low levels of these traits. By contrast, differences in conscientiousness (high vs. low) do not yield statistically significant variations in acceptances or callbacks.

The complementary field settings – searching for (online) friends and (offline) roommates – provide a comprehensive perspective on how online personality signals translate into tangible, informal social outcomes. Overall, our findings offer causal evidence that online impressions can significantly affect both online and offline outcomes and decision-making processes, extending beyond what has been shown in laboratory or vignette-based studies (Stoughton et al., 2013; Utz, 2010; van Zoonen & van der Meer, 2015; Walther et al., 2009; Walther et al., 2008). Future field research could build on these insights by exploring additional personality traits, examining gender differences, or investigating whether conscientiousness becomes more salient in professional or performance-based contexts.

ESG Metrics and Executive Compensation Finally, Chapter 6 shifts from field experiments manipulating social media information to corporate governance, connecting the broader theme of *non-traditional* information to sustainability criteria in executive compensation, beyond standard financial metrics. Although ESG metrics are increasingly included in executive remuneration, many appear to be discretionary, assigned low weights, and thus have limited capacity to affect executive behavior, raising concerns about "greenwashing." In this chapter, my co-authors and I hand-collect extensive data on ex-ante compensation packages, ex-post target realizations, and ultimately realized pay, surpassing the breadth and detail of prior studies that rely on commercially available datasets. Whereas such datasets typically provide only a single indicator of whether any ESG metric was included, our approach captures the scope, structure, weight, *and* realization of these metrics.

The findings show that most ESG metrics contribute little to executive pay risk, with weak incentives especially prevalent in financial firms and large companies – particularly at the most visible executive levels – reinforcing concerns about greenwashing. By contrast, in sectors with substantial environmental footprints, ESG metrics more frequently carry significant payout weights, suggesting that external scrutiny and reputational or regulatory pressures can foster more robust incentive designs.

By demonstrating the frequent use of "weak" ESG incentives, we contribute to the ongoing debate over whether these metrics genuinely align corporate decision-making with broader societal goals or merely serve as a superficial signal of social responsibility. Despite drawing on a detailed panel dataset, potential endogeneity issues remain; for example, firms that truly prioritize sustainability may also be more inclined to adopt robust ESG metrics. Establishing the causal impact of ESG inclusion on executive decision-making thus remains challenging, particularly in the absence of experimental or quasi-experimental variation – an area that future research could address more definitively.

Across the chapters, a common message emerges: merely introducing new information or metrics – on social media or in executive compensation – does not automatically achieve the intended goals of reducing bias, promoting equity, or enhancing sustainability. Context matters, as do the structural forces that shape how decision-makers process these signals. Careful attention to how these signals are constructed, perceived, and enforced is critical.

Nonetheless, the studies also highlight signs of progress. They show that stereotype-breaking information can partially erode persistent discrimination, and that firms in environmentally sensitive sectors are more likely to adopt ESG metrics that actually matters. These findings highlight the potential to harness informational interventions and well-structured metrics to promote fairer, more accountable practices in both interpersonal and corporate settings.

Beyond academic discourse, these findings hold practical relevance. Understanding the subtle ways in which *non-traditional* social media information can ameliorate, but also reinforce stereotypes can guide platform design and anti-discrimination policies. Meanwhile, insights into ESG target's actual effect on executive compensation can inform policymakers, investors, and activists aiming to align corporate behavior with broader societal goals. Ultimately, the introduction of new information or metrics into decision-making processes – whether at the individual, market, or corporate level – does not automatically yield the intended outcomes. Without carefully considered design, oversight, and (if necessary) regulatory frameworks, well-intentioned innovations risk becoming superficial fixes that leave entrenched inequalities and greenwashing intact.

In conclusion, this thesis contributes to our understanding of how *non-traditional* informational, i.e., digital footprints or non-financial performance measures, reshape behavior in informal markets and organizations. By identifying both the promise and pitfalls of these signals, it points toward a continued need for empirical research and evidence-based policies aimed at mitigating bias, fostering inclusion, and ensuring that sustainability goals are met in practice rather than merely in principle.

References

Acquisti, A., & Fong, C. (2020). An experiment in hiring discrimination via online social networks. *Management Science*, *66*(3), 1005–1024. https://doi.org/10.1287/mnsc.2018.3269

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science*, *347*(6221), 509–514. https://doi.org/10.1126/science.aaa1465

Aggarwal, R. K., & Samwick, A. A. (1999). The other side of the trade-off: The impact of risk on executive compensation. *Journal of Political Economy*, *107*(1), 65–105. https://doi.org/10.1086/250051

Ahmed, A. M., Andersson, L., & Hammarstedt, M. (2010). Can discrimination in the housing market be reduced by increasing the information about the applicants? *Land Economics*, *86*(1), 79–90. https://doi.org/10.3368/le.86.1.79

Aigner, D. J., & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *ILR Review*, *30*(2), 175–187. https://doi.org/10.1177/001979397703000204

Ajzenman, N., Ferman, B., & C Sant'Anna, P. (2025). Discrimination in the formation of academic networks: A field experiment on #econtwitter [Forthcoming]. *American Economic Review: Insights*. https://www.aeaweb.org/articles?id=10.1257/aeri.20240298

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, *84*(3), 488–500. https://doi.org/10.2307/1879431

Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2020). The welfare effects of social media. *American Economic Review*, *110*(3), 629–76. https://doi.org/10.1257/aer.20190658

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211–36. https://doi.org/10.1257/jep.31.2.211

Allport, G. W. (1954). *The nature of prejudice*. Addison-Wesley.

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological monographs*, *47*(1), i–171.

Alsan, M., Garrick, O., & Graziani, G. (2019). Does diversity matter for health? experimental evidence from oakland. *American Economic Review*, *109*(12), 4071–4111. https://doi.org/10.1257/aer.20181446

Al-Shaer, H., Albitar, K., & Liu, J. (2023). Ceo power and csr-linked compensation for corporate environmental responsibility: Uk evidence. *Review of Quantitative Finance and Accounting*, *60*(3), 1025–1063.

Al-Shaer, H., & Zaman, M. (2019). Ceo compensation and sustainability reporting assurance: Evidence from the uk. *Journal of Business Ethics*, *158*, 233–252.

Altonji, J. G., & Blank, R. M. (1999). Race and gender in the labor market. Elsevier. https://doi.org/10.1016/S1573-4463(99)30039-0

Altonji, J. G., & Pierret, C. R. (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, *116*(1), 313–350. https://doi.org/10.1162/003355301556329

Aneja, A., Luca, M., & Reshef, O. (2025). The benefits of revealing race: Evidence from minority-owned local businesses. *American Economic Review*, *115*(2), 660–689. https://doi.org/10.1257/aer.20230075

Aresu, S., Hooghiemstra, R., & Melis, A. (2023). Integration of csr criteria into executive compensation contracts: A cross-country analysis. *Journal of Management*, *49*(8), 2766–2804.

Aridor, G., Jiménez-Durán, R., Levy, R., & Song, L. (2024). The economics of social media. *Journal of Economic Literature*, *62*(4), 1422–74. https://doi.org/10.1257/jel.20241743

Arrow, K. J. (1973). The theory of discrimination. *Discrimination in Labor Markets*, *3*(10), 3–33. https://EconPapers.repec.org/RePEc:pri:indrel:30a

Arrow, K. J. (1998). What has economics to say about racial discrimination? *Journal of Economic Perspectives*, *12*(2). https://doi.org/10.1257/jep.12.2.91

Asendorpf, J. B., & Wilpers, S. (1998). Personality effects on social relationships. *Journal of Personality and Social Psychology*, *74*(6), 1531–1544.

Ash, E., Durante, R., Grebenschikova, M., & Schwarz, C. (2021). *Visual representation and stereotypes in news media* [CEPR Discussion Paper No. DP16624, Available at SSRN: https://ssrn.com/abstract=3960205].

Auspurg, K., Schneck, A., & Hinz, T. (2019). Closed doors everywhere? a meta-analysis of field experiments on ethnic discrimination in rental housing markets. *Journal of Ethnic and Migration Studies*, *45*(1), 95–114. https://doi.org/10.1080/1369183X.2018.1489223

Azar, J., Duro, M., Kadach, I., & Ormazabal, G. (2021). The big three and corporate carbon emissions around the world. *Journal of Financial Economics*, *142*(2), 674–696.

Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., & Gosling, S. D. (2010). Facebook profiles reflect actual personality, not self-idealization. *Psychological Science*, *21*(3), 372–374. https://doi.org/10.1177/0956797609360756

Baert, S. (2018). Facebook profile picture appearance affects recruiters' first hiring decisions. *New Media & Society*, *20*(3), 1220–1239. https://doi.org/10.1177/1461444816687294

Baert, S., Albanese, A., du Gardein, S., Ovaere, J., & Stappers, J. (2017). Does work experience mitigate discrimination? *Economics Letters*, *155*, 35–38. https://doi.org/10.1016/j.econlet.2017.03.011

Bailey, J., Steeves, V., Burkell, J., & Regan, P. (2013). Negotiating with gender stereotypes on social networking sites: From "bicycle face" to facebook. *Journal of Communication Inquiry*, *37*(2), 91–112. https://doi.org/10.1177/0196859912473777

Barontini, R., & Hill, J. G. (2024). Sustainability and executive compensation. In K. Alexander, M. Gargantini, & M. Siri (Eds.), *The cambridge handbook of eu sustainable finance: Regulation, supervision, and governance*. Cambridge University Press.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*(1), 1–26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, *9*(1-2), 9–30. https://doi.org/10.1111/1468-2389.00160

Bartoš, V., Bauer, M., Chytilová, J., & Matějka, F. (2016). Attention discrimination: Theory and field experiments with monitoring information acquisition. *American Economic Review*, *106*(6), 1437–75. https://doi.org/10.1257/aer.20140571

Barzuza, M., Curtis, Q., & Webber, D. H. (2023). The millennial corporation: Strong stakeholders, weak managers. *Stan. JL Bus. & Fin.*, *28*, 255–306.

Baur, R., & Ossenberg, S. (2016). Deutsche und türkische Stereotype im Vergleich [https://www.uni-due.de/imperia/md/images/ikk/ergebnisse_side.pdf, Retrieved: September 27, 2023].

Bayer, P., & Charles, K. K. (2018). Divergent paths: A new perspective on earnings differences between black and white men since 1940. *The Quarterly Journal of Economics*, *133*(3), 1459–1501. https://doi.org/10.1093/qje/qjy003

Bayer, P., Ferreira, F., & Ross, S. L. (2016). The vulnerability of minority homeowners in the housing boom and bust. *American Economic Journal: Economic Policy*, *8*(1), 1–27. https://doi.org/10.1257/pol.20140074

Beaman, L., Duflo, E., Pande, R., & Topalova, P. (2012). Female leadership raises aspirations and educational attainment for girls: A policy experiment in india. *Science*, *335*(6068), 582–586. https://doi.org/10.1126/science.1212382

Bebchuk, L. A., & Tallarita, R. (2022). The perils and questionable promise of esg-based compensation. *J. Corp. L.*, *48*, 37.

Becker, G. S. (1957). *The economics of discrimination*. University of Chicago Press.

Becker, J. C., & Swim, J. K. (2011). Seeing the unseen: Attention to daily encounters with sexism as way to reduce sexist beliefs. *Psychology of Women Quarterly*, *35*(2), 227–242. https://doi.org/10.1177/0361684310397509

Becton, J. B., Walker, H. J., Schwager, P., & Gilstrap, J. B. (2019). Is what you see what you get? investigating the relationship between social media content and counterproductive work behaviors, alcohol consumption, and episodic heavy drinking. *The International Journal of Human Resource Management*, *30*(15), 2251–2272. https://doi.org/10.1080/09585192.2017.1314977

Belleflamme, P., & Peitz, M. (2021). *The economics of platforms*. Cambridge University Press. https://doi.org/10.1017/9781108696913

Benjamin, D. J., Choi, J. J., & Strickland, A. J. (2010). Social identity and preferences. *American Economic Review*, *100*(4), 1913–28. https://doi.org/10.1257/aer.100.4.1913

Berg, F., Fabisik, K., & Sautner, Z. (2020). Rewriting history ii: The (un) predictable past of esg ratings. *European Corporate Governance Institute–Finance Working Paper*, *708*(2020), 10–2139.

Bertrand, M., & Duflo, E. (2017). Field experiments on discrimination. In A. V. Banerjee & E. Duflo (Eds.), *Handbook of field experiments* (pp. 309–393). North-Holland. https://doi.org/10.1016/bs.hefe.2016.08.004

Bertrand, M., & Mullainathan, S. (2004). Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American Economic Review*, *94*(4), 991–1013. https://doi.org/10.1257/0002828042002561

Bohren, J. A., Haggag, K., Imas, A., & Pope, D. G. (2023). Inaccurate statistical discrimination: An identification problem. *The Review of Economics and Statistics*, 1–45. https://doi.org/10.1162/rest_a_01367

Bohren, J. A., Imas, A., & Rosenberg, M. (2019). The dynamics of discrimination: Theory and evidence. *American Economic Review*, *109*(10), 3395–3436. https://doi.org/10.1257/aer.20171829

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, *489*(7415), 295–298. https://doi.org/10.1038/nature11421

Bonham, J., & Riggs-Cragun, A. (2024). Motivating esg activities through contracts, taxes and disclosure regulation. *Chicago Booth Research Paper*, (22-05).

Bordalo, P., Coffman, K., Gennaioli, N., & Shleifer, A. (2016). Stereotypes. *The Quarterly Journal of Economics*, *131*(4), 1753–1794. https://doi.org/10.1093/qje/qjw029

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly Journal of Economics*, *127*(3), 1243–1285. https://doi.org/10.1093/qje/qjs018

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, *121*(5), 803–843.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2022). Salience. *Annual Review of Economics*, *14*, 521–544. https://doi.org/10.1146/annurev-economics-051520-011616

Bordalo, P., Tabellini, M., & Yang, D. Y. (2020). *Issue salience and political stereotypes* (Working Paper No. 27194). National Bureau of Economic Research. https://doi.org/10.3386/w27194

Bosch, M., Carnero, M. A., & Farré, L. (2010). Information and discrimination in the rental housing market: Evidence from a field experiment. *Regional Science and Urban Economics*, *40*(1), 11–19. https://doi.org/10.1016/j.regsciurbeco.2009.11.001

Brickley, J. A., Smith, C. W., & Zimmerman, J. L. (2017). *Managerial economics and organizational architecture*. McGraw-Hill Education.

Broccardo, E., Hart, O., & Zingales, L. (2022). Exit versus voice. *Journal of Political Economy*, *130*(12), 3101–3145. https://doi.org/10.1086/720516

Brown, V. R., & Vaughn, E. (2011). The writing on the (facebook) wall: The use of social networking sites in hiring decisions. *Journal of Business and Psychology*, *26*(2), 219–225. https://doi.org/10.1007/s10869-011-9221-x

Bursell, M. (2007). What's in a name? - a field experiment test for the existence of ethnic discrimination in the hiring process [Retrieved: February 8, 2024]. https://www.diva-portal.org/smash/get/diva2:821429/FULLTEXT01.pdf

Buss, D. M. (1991). Evolutionary personality psychology. *Annual Review of Psychology*, *42*(1), 459–491.

Campos-Mercade, P., & Mengel, F. (2023). Non-bayesian statistical discrimination. *Management Science*. https://doi.org/10.1287/mnsc.2023.4824

Carley, K. (1993). Coding choices for textual analysis: A comparison of content analysis and map analysis. *Sociological Methodology*, *23*, 75–126.

Carlsson, M., & Eriksson, S. (2014). Discrimination in the rental market for apartments. *Journal of Housing Economics*, *23*, 41–54. https://doi.org/10.1016/j.jhe.2013.11.004

Carlsson, M., & Eriksson, S. (2015). Ethnic discrimination in the london market for shared housing. *Journal of Ethnic and Migration Studies*, *41*(8), 1276–1301. https://doi.org/10.1080/1369183X.2014.965670

Carlsson, M., & Rooth, D.-O. (2012). Revealing taste-based discrimination in hiring: A correspondence testing experiment with geographic variation. *Applied Economics Letters*, *19*(18), 1861–1864. https://doi.org/10.1080/13504851.2012.667537

Carpusor, A. G., & Loges, W. E. (2006). Rental discrimination and ethnicity in names. *Journal of Applied Social Psychology*, *36*(4), 934–952. https://doi.org/10.1111/j.0021-9029.2006.00050.x

Carter, M. E., Pawliczek, A., & Zhong, R. I. (2023). Say on esg: The adoption of say-on-pay laws, esg contracting, and firm esg performance. *European Corporate Governance Institute–Finance Working Paper*, (886).

Cheng, H. K., Daniel Sokol, D., & Zang, X. (2024). The rise of empirical online platform research in the new millennium. *Journal of Economics & Management Strategy*, *33*(2), 416–451. https://doi.org/10.1111/jems.12571

Chetty, R., Hendren, N., Jones, M. R., & Porter, S. R. (2020). Race and economic opportunity in the united states: An intergenerational perspective*. *The Quarterly Journal of Economics*, *135*(2), 711–783. https://doi.org/10.1093/qje/qjz042

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states *. *The Quarterly Journal of Economics*, *129*(4), 1553–1623. https://doi.org/10.1093/qje/qju022

Chetty, R., Looney, A., & Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, *99*(4), 1145–77. https://doi.org/10.1257/aer.99.4.1145

Cho, C. H., & Roberts, R. W. (2010). Environmental reporting on the internet by america's toxic 100: Legitimacy and self-presentation. *International Journal of Accounting Information Systems*, *11*(1), 1–16.

Christensen, H. B., Hail, L., & Leuz, C. (2018). Economic analysis of widespread adoption of csr and sustainability reporting standards. *Available at SSRN 3315673*.

Clark, V., Tuffin, K., Frewin, K., & Bowker, N. (2018). Housemate desirability and understanding the social dynamics of shared living. *Qualitative Psychology*, *5*(1), 26. https://doi.org/10.1037/qup0000091

Coffman, K., Collis, M. R., & Kulkarni, L. (2023). Stereotypes and Belief Updating. *Journal of the European Economic Association*, *22*(3), 1011–1054. https://doi.org/10.1093/jeea/jvad063

Cohen, S., Kadach, I., & Ormazabal, G. (2023). Institutional investors, climate disclosure, and carbon emissions. *Journal of Accounting and Economics*, *76*(2-3), 101640.

Cohen, S., Kadach, I., Ormazabal, G., & Reichelstein, S. (2023). Executive compensation tied to esg performance: International evidence. *Journal of Accounting Research*, *61*(3), 805–853.

Cornell, B., & Welch, I. (1996). Culture, information, and screening discrimination. *Journal of political Economy*, *104*(3), 542–571. https://doi.org/10.1086/262033

Correll, J., Wittenbrink, B., Crawford, M. T., & Sadler, M. S. (2015). Stereotypic vision: How stereotypes disambiguate visual stimuli. *Journal of Personality and Social Psychology*, *108*(2), 219–233.

Crilly, D., Hansen, M., & Zollo, M. (2016). The grammar of decoupling: A cognitive-linguistic perspective on firms' sustainability claims and stakeholders' interpretation. *Academy of Management Journal*, *59*(2), 705–729.

Cucurull, G., Rodríguez, P., Yazici, V. O., Gonfaus, J. M., Roca, F. X., & Gonzàlez, J. (2018). Deep inference of personality traits by integrating image and word use in social networks. https://arxiv.org/abs/1802.06757

Cui, R., Li, J., & Zhang, D. J. (2020). Reducing discrimination with reviews in the sharing economy: Evidence from field experiments on airbnb. *Management Science*, *66*(3), 1071–1094. https://doi.org/10.1287/mnsc.2018.3273

Datareportal. (2021). *Digital 2021: Global overview report* [Retrieved: May 29, 2022]. https://datareportal.com/reports/digital-2021-global-overview-report

Datareportal. (2024). Instagram users, stats, data & trends [Retrieved: November 28, 2024].

de la Llama, V. A., Trueba, I., Voges, C., Barreto, C., & Park, D. J. (2012). At face(book) value: Uses of facebook in hiring processes and the role of identity in social networks. *International Journal of Work Innovation*, *1*(1), 114–136. https://doi.org/10.1504/IJWI.2012.047984

De Leyn, T. (2023). Reclaiming agency in the digital neighborhood: an ethnographic exploration of ethno-religious minority youths' performances of the masculine self. *Journal of Computer-Mediated Communication*, *28*(6), zmad037. https://doi.org/10.1093/jcmc/zmad037

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature*, *47*(2), 315–72. https://doi.org/10.1257/jel.47.2.315

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of personality and social psychology*, *56*(1), 5–18.

Diehl, C., Andorfer, V. A., Khoudja, Y., & Krause, K. (2013). Not in my kitchen? ethnic discrimination and discrimination intentions in shared housing among university students in germany. *Journal of Ethnic and Migration Studies*, *39*(10), 1679–1697. https://doi.org/10.1080/1369183X.2013.833705

Doleac, J. L., & Stein, L. C. (2013). The visible hand: Race and online market outcomes. *The Economic Journal*, *123*(572), F469–F492. https://doi.org/10.1111/ecoj.12082

Domahidi, E., Merkt, M., Thiersch, C., Utz, S., & Schüler, A. (2022). You want this job? influence and interplay of self-generated text and picture cues in professional networking service profiles on expertise evaluation. *Media Psychology*, *25*(2), 290–317. https://doi.org/10.1080/15213269.2021.1927104

Donohue III, J. J., & Levitt, S. D. (2001). The impact of race on policing and arrests. *The Journal of Law and Economics*, *44*(2), 367–394. https://doi.org/10.1086/322810

Dustmann, C., Glitz, A., Schönberg, U., & Brücker, H. (2015). Referral-based Job Search Networks. *The Review of Economic Studies*, *83*(2), 514–546. https://doi.org/10.1093/restud/rdv045

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science*, *60*(11), 2835–2857. https://doi.org/10.1287/mnsc.2014.1984

Edelman, B., Luca, M., & Svirsky, D. (2017). Racial discrimination in the sharing economy: Evidence from a field experiment. *American Economic Journal: Applied Economics*, *9*(2), 1–22. https://doi.org/10.1257/app.20160213

Edmans, A. (2011). Does the stock market fully value intangibles? employee satisfaction and equity prices. *Journal of Financial Economics*, *101*(3), 621–640.

Edmans, A. (2023). The end of esg. *Financial Management*, *52*(1), 3–17.

Edmans, A. (2024). Rational sustainability. *Journal of Applied Corporate Finance*.

Edmans, A., Gabaix, X., & Jenter, D. (2017). Executive compensation: A survey of theory and evidence. In B. E. Hermalin & M. S. Weisbach (Eds.), *The handbook of the economics of corporate governance* (pp. 383–539). North-Holland. https://doi.org/10.1016/bs.hecg.2017.11.010

Edmans, A., Gosling, T., & Jenter, D. (2023). Ceo compensation: Evidence from the field. *Journal of Financial Economics*, *150*(3), 103718.

Ekman, M. (2019). Anti-immigration and racist discourse in social media. *European Journal of Communication*, *34*(6), 606–618. https://doi.org/10.1177/0267323119886151

Evsyukova, Y., Rusche, F., & Mill, W. (2025). Linkedout? a field experiment on discrimination in job network formation. *The Quarterly Journal of Economics*, *140*(1), 283–334. https://doi.org/10.1093/qje/qjae035

Fang, H., & Moro, A. (2011). Theories of statistical discrimination and affirmative action: A survey. In J. Benhabib, A. Bisin, & M. O. Jackson (Eds.). North-Holland. https://doi.org/10.1016/B978-0-444-53187-2.00005-X

Fang, R., Landis, B., Zhang, Z., Anderson, M. H., Shaw, J. D., & Kilduff, M. (2015). Integrating personality and social networks: A meta-analysis of personality, network position, and work outcomes in organizations. *Organization Science*, *26*(4), 1243–1260. https://doi.org/10.1287/orsc.2015.0972

Federal Statistical Office. (2022a). *Ausländer: Deutschland, stichtag, geschlecht/altersjahre/familienstand* [Retrieved: August 1, 2022]. https://www-genesis.destatis.de/genesis//online?operation=table&code=12521-0001

Federal Statistical Office. (2022b). *Migration und integration: Bevölkerung in privathaushalten nach migrationshintergrund* [Retrieved: August 1, 2022]. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Tabellen/migrationshintergrund-geschlecht-insgesamt.html

Federal Statistical Office. (2022c). *Migration und integration: Bevölkerung in privathaushalten nach migrationshintergrund* [Retrieved: August 1, 2022]. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Tabellen/migrationshintergrund-staatsangehoerigkeit-staaten.html

Federal Statistical Office. (2022d). *Studierende: Deutschland, semester, nationalität, geschlecht, studienfach* [Retrieved: August 1, 2022]. https://www-genesis.destatis.de/genesis//online?operation=table&code=21311-0003

Federal Statistical Office. (2024a). Migration und integration – bevölkerung nach migrationshintergrund [Retrieved: February 4, 2025].

Federal Statistical Office. (2024b). Migration und integration – bevölkerung nach migrationshintergrund nach ausgewählten geburtsstaaten [Retrieved: February 4, 2025].

Flammer, C., Hong, B., & Minor, D. (2019). Corporate governance and the rise of integrating corporate social responsibility criteria in executive compensation: Effectiveness and implications for firm outcomes. *Strategic Management Journal*, *40*(7), 1097–1122.

Frison, E., & Eggermont, S. (2017). Browsing, posting, and liking on instagram: The reciprocal relationships between different types of instagram use and adolescents' depressed mood. *Cyberpsychology, Behavior, and Social Networking*, *20*(10), 603–609. https://doi.org/10.1089/cyber.2017.0156

Frydman, C., & Wang, B. (2020). The impact of salience on investor behavior: Evidence from a natural experiment. *The Journal of Finance*, *75*(1), 229–276. https://doi.org/10.1111/jofi.12851

Fryer, R. G. (2019). An empirical analysis of racial differences in police use of force. *Journal of Political Economy*, *127*(3), 1210–1261. https://doi.org/10.1086/701423

Gabaix, X. (2014). A sparsity-based model of bounded rationality. *The Quarterly Journal of Economics*, *129*(4), 1661–1710. https://doi.org/10.1093/qje/qju024

Galos, D. R. (2023). Social media and hiring: a survey experiment on discrimination based on online social class cues. *European Sociological Review*, jcad012. https://doi.org/10.1093/esr/jcad012

Garel, A., Romec, A., Sautner, Z., & Wagner, A. F. (2024). Do investors care about biodiversity? *Review of Finance*, *28*(4), 1151–1186.

Garen, J. E. (1994). Executive compensation and principal-agent theory. *Journal of Political Economy*, *102*(6), 1175–1199. https://doi.org/10.1086/261967

Garvey, G., & Milbourn, T. (2003). Incentive compensation when executives can hedge the market: Evidence of relative performance evaluation in the cross section. *The Journal of Finance*, *58*(4), 1557–1582. https://doi.org/10.1111/1540-6261.00577

GEHRING, K., HARM ADEMA, J. A., & POUTVAARA, P. (2022). *Immigrant narratives* [CESifo Working Paper No. 10026, Available at SSRN: https://ssrn.com/abstract=4259623].

GESELLSCHAFT FÜR DEUTSCHE SPRACHE. (2019). Vornamen. [Retrieved: July 3, 2019]. https://gfds.de/vornamen

GHOSE, A., IPEIROTIS, P. G., & LI, B. (2014). Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science*, *60*(7), 1632–1654. https://doi.org/10.1287/mnsc.2013.1828

GOSLING, S. D., GADDIS, D., & VAZIRE, S. (2007). Personality Impressions based on Facebook profiles. *ICWSM*. https://icwsm.org/papers/3--Gosling-Gaddis-Vazire.pdf

GRABER, D. A. (1990). Seeing is remembering: How visuals contribute to learning from television news. *Journal of Communication*, *40*(3), 134–155. https://doi.org/10.1111/j.1460-2466.1990.tb02275.x

GRABNER, I., RENDERS, A., & YANG, L. (2024). The complementarity between csr disclosures and the use of csr-based performance measures in ceo annual incentive contracts. *Available at SSRN 3670738*.

GUESS, A. M., MALHOTRA, N., PAN, J., BARBERÁ, P., ALLCOTT, H., BROWN, T., CRESPO-TENORIO, A., DIMMERY, D., FREELON, D., GENTZKOW, M., GONZÁLEZ-BAILÓN, S., KENNEDY, E., KIM, Y. M., LAZER, D., MOEHLER, D., NYHAN, B., RIVERA, C. V., SETTLE, J., THOMAS, D. R., … TUCKER, J. A. (2023). How do social media feed algorithms affect attitudes and behavior in an election campaign? *Science*, *381*(6656), 398–404. https://doi.org/10.1126/science.abp9364

GUNARATHNE, P., RUI, H., & SEIDMANN, A. (2022). Racial bias in customer service: Evidence from twitter. *Information Systems Research*, *33*(1), 43–54. https://doi.org/10.1287/isre.2021.1058

HALL, J. A., PENNINGTON, N., & LUEDERS, A. (2014). Impression management and formation on Facebook: A lens model approach. *New Media & Society*, *16*(6), 958–982. https://doi.org/10.1177/1461444813495166

HAMILTON, D. L., FALLOT, R. D., & HAUTALUOMA, J. (1978). Information salience and order effects in impression formation. *Personality and Social Psychology Bulletin*, *4*(1), 44–47. https://doi.org/10.1177/014616727800400108

HARRIS, K., & VAZIRE, S. (2016). On friendship development and the Big Five personality traits. *Social and Personality Psychology Compass*, *10*(11), 647–667. https://doi.org/10.1111/spc3.12287

HARRISON, G. W., & LIST, J. A. (2004). Field experiments. *Journal of Economic Literature*, *42*(4), 1009–1055. https://doi.org/10.1257/0022051043004577

HART, O., & ZINGALES, L. (2022). The new corporate governance. *The University of Chicago Business Law Review*, *1*, 195–216.

HARTMANN, S., & KNAPPE, A. (2019). *Der studentische wohnungsmarkt 2019: Analysen zur bundesweiten und standortspezifischen wohnsituation studierender in deutschland* [Retrieved: May 29, 2022]. https://www.constata.de/downloads/2019_Studie%20Fachkraft2030_Der%20studentische%20Wohnungsmarkt%202019.pdf

Hazarika, S., Kashikar, A., Peng, L., Röell, A., & Shen, Y. (2022). Esg-linked pay around the world—trends, determinants, and outcomes. *Working Paper*.

Heckman, J. J. (1998). Detecting discrimination. *Journal of Economic Perspectives*, *12*(2), 101–116. https://doi.org/10.1257/jep.12.2.101

Hermes, H., Lergetporer, P., Mierisch, F., Peter, F., & Wiederhold, S. (2024). *Discrimination in universal social programs? a nationwide field experiment on access to child care* (IWH Discussion Papers No. 12/2023). Halle (Saale), Halle Institute for Economic Research (IWH). https://hdl.handle.net/10419/284361

Hilton, J. L., & von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology*, *47*, 237–271. https://doi.org/10.1146/annurev.psych.47.1.237

Himmelberg, C. P., Hubbard, R., & Palia, D. (1999). Understanding the determinants of managerial ownership and the link between ownership and performance. *Journal of Financial Economics*, *53*(3), 353–384. https://doi.org/10.1016/S0304-405X(99)00025-2

Holmström, B. (1979). Moral hazard and observability. *The Bell Journal of Economics*, 74–91.

Homroy, S., Mavruk, T., & Nguyen, V. D. (2023). Esg-linked compensation, ceo skills, and shareholder welfare. *The Review of Corporate Finance Studies*, *12*(4), 939–985.

Hong, B., Li, Z., & Minor, D. (2016). Corporate governance and executive compensation for corporate social responsibility. *Journal of Business Ethics*, *136*, 199–213.

Hu, Y., Manikonda, L., & Kambhampati, S. (2014). What we instagram: A first analysis of instagram photo content and user types. *Eighth International AAAI Conference on Weblogs and Social Media*, 595–598.

Ikram, A., Li, Z. F., & Minor, D. (2023). Csr-contingent executive compensation contracts. *Journal of Banking & Finance*, *151*, 105655.

Insider Intelligence. (2022). Instagram user stats, demographics & trends in 2022 [Retrieved: February 8, 2024]. https://www.insiderintelligence.com/chart/c/T11780/instagram-users-worldwide-2020-2025-millions-change-of-internet-users-1

Ioannides, Y. M., & Datcher Loury, L. (2004). Job information networks, neighborhood effects, and inequality. *Journal of Economic Literature*, *42*(4), 1056–1093. https://doi.org/10.1257/0022051043004595

Jackson, K. M., & Trochim, W. M. K. (2002). Concept mapping as an alternative approach for the analysis of open-ended survey responses. *Organizational Research Methods*, *5*(4), 307–336. https://doi.org/10.1177/109442802237114

Jensen-Campbell, L. A., Adams, R., Perry, D. G., Workman, K. A., Furdella, J. Q., & Egan, S. K. (2002). Agreeableness, extraversion, and peer relations in early adolescence: Winning friends and deflecting aggression. *Journal of Research in Personality*, *36*(3), 224–251.

Jensen-Campbell, L. A., & Malcolm, K. T. (2007). The importance of conscientiousness in adolescent interpersonal relationships. *Personality and Social Psychology Bulletin*, *33*(3), 368–383. https://doi.org/10.1177/0146167206296104

Jin, L. (2002). Ceo compensation, diversification, and incentives. *Journal of Financial Economics*, *66*(1), 29–63. https://doi.org/10.1016/S0304-405X(02)00150-2

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big five trait taxonomy. *Handbook of personality: Theory and research*, *3*(2), 114–158.

Judge, T. A., Higgins, C. A., Thoresen, C. J., & Barrick, M. R. (1999). The big five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, *52*(3), 621–652. https://doi.org/10.1111/j.1744-6570.1999.tb00174.x

Kaas, L., & Manger, C. (2012). Ethnic discrimination in germany's labour market: A field experiment. *German Economic Review*, *13*(1), 1–20. https://doi.org/doi:10.1111/j.1468-0475.2011.00538.x

Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate sustainability: First evidence on materiality. *The Accounting Review*, *91*(6), 1697–1724.

Kluemper, D. H., & Rosen, P. A. (2009). Future employment selection methods: Evaluating social networking web sites. *Journal of Managerial Psychology*, *24*(6), 567–580. https://doi.org/10.1108/02683940910974134

Kluemper, D. H., Rosen, P. A., & Mossholder, K. W. (2012). Social networking websites, personality ratings, and the organizational context: More than meets the eye? *Journal of Applied Social Psychology*, *42*(5), 1143–1172. https://doi.org/10.1111/j.1559-1816.2011.00881.x

Koch, W. (2023). Ergebnisse der ard/zdf-onlinestudie 2023: Soziale medien werden 30 minuten am tag genutzt – instagram ist die plattform nummer eins [Retrieved: January 20, 2024].

Kohlheim, R., & Kohlheim, V. (2005). *Duden – familiennamen: Herkunft und bedeutung von 20.000 nachnamen* (Vol. 2. Auflage). Bibliographisches Institut, Berlin; Duden Verlag.

Kolk, A., & Perego, P. (2014). Sustainable bonuses: Sign of corporate responsibility or window dressing? *Journal of Business Ethics*, *119*, 1–15.

Kroher, M., Beusse, M., Isleib, S., Becker, K., Ehrhardt, M.-C., Gerdes, F., Koopmann, J., Schommer, T., Schwabe, U., Steinkühler, J., Völk, D., Peter, F., & Buchholz, S. (2023). Die studierendenbefragung in deutschland: 22. sozialerhebung [Retrieved: October 18, 2024].

Krueger, P., Sautner, Z., & Starks, L. T. (2020). The importance of climate risks for institutional investors. *The Review of Financial Studies*, *33*(3), 1067–1111.

Lambert, R. A., & Larcker, D. F. (1987). An analysis of the use of accounting and market measures of performance in executive compensation contracts. *Journal of Accounting Research*, 85–125.

Lambrecht, A., & Tucker, C. (2019). Algorithmic bias? an empirical study of apparent gender-based discrimination in the display of stem career ads. *Management Science*, *65*(7), 2966–2981. https://doi.org/10.1287/mnsc.2018.3093

Landier, A., & Lovo, S. (2024). Socially responsible finance: How to optimize impact. *The Review of Financial Studies*, hhae055. https://doi.org/10.1093/rfs/hhae055

Lang, K., & Kahn-Lang Spitzer, A. (2020). Race discrimination: An economic perspective. *Journal of Economic Perspectives*, *34*(2), 68–89. https://doi.org/10.1257/jep.34.2.68

Lee, S. Y., Pitesa, M., Thau, S., & Pillutla, M. M. (2015). Discrimination in selection decisions: Integrating stereotype fit and interdependence theories. *Academy of Management Journal*, *58*(3), 789–812. https://doi.org/10.5465/amj.2013.0571

LEINER, D. J. (2019). Too fast, too straight, too weird: Non-reactive indicators for meaningless data in internet surveys. *Survey Research Methods*, *13*(3), 229–248. https://doi.org/10.18148/srm/2019.v13i3.7403

LEVY, R. (2021). Social media, news consumption, and polarization: Evidence from a field experiment. *American Economic Review*, *111*(3), 831–70. https://doi.org/10.1257/aer.20191777

LI, Z., & WANG, L. (2016). Executive compensation incentives contingent on long-term accounting performance. *The Review of Financial Studies*, *29*(6), 1586–1633.

LIPPENS, L., VERMEIREN, S., & BAERT, S. (2023). The state of hiring discrimination: A meta-analysis of (almost) all recent correspondence experiments. *European Economic Review*, *151*, 104315. https://doi.org/10.1016/j.euroecorev.2022.104315

LSEG. (2024). Environmental, social and governance scores from lseg [https://www.lseg.com/content/dam/data-analytics/en_us/documents/methodology/lseg-esg-scores-methodology.pdf [Accessed: March 03, 2024]].

LUCA, M. (2016). *Reviews, reputation, and revenue: The case of yelp.com* (tech. rep. No. 12-016) [Retrieved: March 2, 2025]. Harvard Business School. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1928601

MAAS, K. (2018). Do corporate social performance targets in executive compensation contribute to corporate social performance? *Journal of Business Ethics*, *148*, 573–585.

MANAGO, A. M., TAYLOR, T., & GREENFIELD, P. M. (2012). Me and my 400 friends: The anatomy of college students' facebook networks, their communication patterns, and well-being. *Developmental psychology*, *48*(2), 369. https://doi.org/10.1037/a0026338

MANANT, M., PAJAK, S., & SOULIÉ, N. (2019). Can social media lead to labor market discrimination? evidence from a field experiment. *Journal of Economics & Management Strategy*, *28*(2), 225–246. https://doi.org/10.1111/jems.12291

MCCRAE, R. R., & COSTA, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of personality and social psychology*, *52*(1), 81.

MCCRAE, R. R., & JOHN, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60*(2), 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

MICHELS, L., OCHMANN, J., SCHMITT, K., LAUMER, S., & TIEFENBECK, V. (2024). Salience, transparency, and self-nudging: A digital nudge to promote healthier food product choices. *European Journal of Information Systems*, *33*(5), 717–747. https://doi.org/10.1080/0960085X.2023.2229787

MIDDENDORFF, E., APOLINARSKI, B., BECKER, K., BORNKESSEL, P., BRANDT, T., HEISSENBERG, S., & POSKOWSKY, J. (2017). *Die wirtschaftliche und soziale lage der studierenden in deutschland 2016* [Retrieved: May 29, 2022]. https://www.bmbf.de/SharedDocs/Publikationen/de/bmbf/4/31338_21_Sozialerhebung_2016_Zusammenfassung.pdf

MLP FINANZBERATUNG SE. (2021). *Mlp studentenwohnreport 2021* [Retrieved: August 1, 2022]. https://mlp-se.de/redaktion/mlp-se-de/studentenwohnreport-microsite/2021/report/mlp-studentenwohnreport-2021-klein.pdf

Moritz, R., & Manger, C. (2022). Somebody that i want to know: The non-monotonic effect of personality information on ethnic and gender discrimination in the market for shared housing. *Journal of Housing Economics*, *57*, 101842. https://doi.org/10.1016/j.jhe.2022.101842

Moritz, R., Manger, C., & Pull, K. (2023). # inviteme: Can social media information reduce discrimination? evidence from a field experiment. *Journal of Economic Behavior & Organization*, *213*, 373–393.

Morse, A., Nanda, V., & Seru, A. (2011). Are incentive contracts rigged by powerful ceos? *The Journal of Finance*, *66*(5), 1779–1821.

Munnell, A. H., Tootell, G. M. B., Browne, L. E., & McEneaney, J. (1996). Mortgage lending in boston: Interpreting hmda data. *The American Economic Review*, *86*(1), 25–53.

Murphy, K. J. (1999). Executive compensation. *Handbook of Labor Economics*, *3*, 2485–2563.

Neal, D. A., & Johnson, W. R. (1996). The role of premarket factors in black-white wage differences. *Journal of Political Economy*, *104*(5), 869–895. https://doi.org/10.1086/262045

Neumark, D. (2012). Detecting discrimination in audit and correspondence studies. *Journal of Human Resources*, *47*(4), 1128–1157. https://doi.org/10.3368/jhr.47.4.1128

Neumark, D. (2018). Experimental research on labor market discrimination. *Journal of Economic Literature*, *56*(3), 799–866. https://doi.org/10.1257/jel.20161309

Nunley, J. M., Owens, M. F., & Howard, R. S. (2011). The effects of information and competition on racial discrimination: Evidence from a field experiment. *Journal of Economic Behavior & Organization*, *80*(3), 670–679. https://doi.org/10.1016/j.jebo.2011.06.028

Nüss, P. (2024). *Management opposition, strikes and union threat* [Working Paper, Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4560917].

Ossenberg, S. (2019). *Deutsche und türkische stereotype: Ein inter-und intrakultureller vergleich*. Springer VS. https://doi.org/10.1007/978-3-658-26879-4

Pager, D. (2007). The use of field experiments for studies of employment discrimination: Contributions, critiques, and directions for the future. *The Annals of the American Academy of Political and Social Science*, *609*(1), 104–133. https://doi.org/10.1177/0002716206294796

Pager, D., & Karafin, D. (2009). Bayesian bigot? statistical discrimination, stereotypes, and employer decision making. *The ANNALS of the American Academy of Political and Social Science*, *621*(1), 70–93. https://doi.org/10.1177/0002716208324628

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*(5), 751–783. https://doi.org/10.1037/0022-3514.90.5.751

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The American Economic Review*, *62*(4), 659–661. http://www.jstor.org/stable/1806107

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*(2), 322–338.

PricewaterhouseCoopers, L. B. S.
bibinitperiod C. f. C. G. (2021). Paying well by paying for good [www.pwc.co.uk/human-resourceservices / assets / pdfs / environmental - social - governance - exec - pay - report . pdf [Accessed: July 23, 2024]].

Qin, B., & Yang, L. (2022). Csr contracting and performance-induced ceo turnover. *Journal of Corporate Finance*, *73*, 102173.

Rammstedt, B., Danner, D., Soto, C. J., & John, O. P. (2020). Validation of the short and extra-short forms of the big five inventory-2 (bfi-2) and their german adaptations. *European Journal of Psychological Assessment*, *36*(1), 149. https://doi.org/10.1027/1015-5759/a000481

Riach, P. A., & Rich, J. (2004). Deceptive field experiments of discrimination: Are they ethical? *Kyklos*, *57*(3), 457–470. https://doi.org/10.1111/j.0023-5962.2004.00262.x

Rosen, P. A., Solomon, S. J., McLarty, B. D., Esken, C. A., & Taylor, E. C. (2018). The use of twitter profiles to assess personality and hireability. *Journal of Managerial Issues*, *30*(2), 256–272. http://www.jstor.org/stable/45176582

Roth, P. L., Bobko, P., Iddekinge, C. H. V., & Thatcher, J. B. (2016). Social media in employee-selection-related decisions: A research agenda for uncharted territory. *Journal of Management*, *42*(1), 269–298. https://doi.org/10.1177/0149206313503018

Rothmann, S., & Coetzer, E. (2003). The big five personality dimensions and job performance. *SA Journal of Industrial Psychology*, *29*(1), 68–74. https://doi.org/10.10520/EJC88938

Sawert, T. (2020). Understanding the mechanisms of ethnic discrimination: A field experiment on discrimination against turks, syrians and americans in the berlin shared housing market. *Journal of Ethnic and Migration Studies*, *46*(19), 3937–3954. https://doi.org/10.1080/1369183X.2019.1577727

Scott, G. G. (2014). More than friends: Popularity on facebook and its role in impression formation. *Journal of Computer-Mediated Communication*, *19*(3), 358–372. https://doi.org/10.1111/jcc4.12067

Segalin, C., Celli, F., Polonio, L., Kosinski, M., Stillwell, D., Sebe, N., Cristani, M., & Lepri, B. (2017). What your facebook profile picture reveals about your personality. *Proceedings of the 25th ACM international conference on Multimedia*, 460–468. https://doi.org/10.1145/3123266.3123331

Selfhout, M., Burk, W., Branje, S., Denissen, J., Van Aken, M., & Meeus, W. (2010). Emerging late adolescent friendship networks and big five personality traits: A social network approach. *Journal of Personality*, *78*(2), 509–538. https://doi.org/https://doi.org/10.1111/j.1467-6494.2010.00625.x

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99–118. https://doi.org/10.2307/1884852

Singh, V. K., Chayko, M., Inamdar, R., & Floegel, D. (2020). Female librarians and male computer programmers? gender bias in occupational images on digital media platforms. *Journal of the Association for Information Science and Technology*, *71*(11), 1281–1294. https://doi.org/10.1002/asi.24335

Soto, C. J., & John, O. P. (2017). Short and extra-short forms of the big five inventory–2: The bfi-2-s and bfi-2-xs. *Journal of Research in Personality*, *68*, 69–81. https://doi.org/j.jrp.2017.02.004

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374. https://doi.org/10.2307/1882010

Spence, M. (2002). Signaling in retrospect and the informational structure of markets. *American Economic Review*, *92*(3), 434–459. https://doi.org/10.1257/00028280260136200

Stopfer, J. M., Egloff, B., Nestler, S., & Back, M. D. (2014). Personality expression and impression formation in online social networks: An integrative approach to understanding the processes of accuracy, impression management and meta–accuracy. *European Journal of Personality*, *28*(1), 73–94. https://doi.org/10.1002/per.1935

Stoughton, J. W., Thompson, L. F., & Meade, A. W. (2013). Big five personality traits reflected in job applicants' social media postings. *Cyberpsychology, Behavior, and Social Networking*, *16*(11), 800–805. https://doi.org/10.1089/cyber.2012.0163

Taylor, S. E., & Thompson, S. C. (1982). Stalking the elusive "vividness" effect. *Psychological Review*, *89*(2), 155–181. https://doi.org/10.1037/0033-295X.89.2.155

Thränhardt, D. (1995). Germany: An undeclared immigration country. *Journal of Ethnic and Migration Studies*, *21*(1), 19–35. https://doi.org/10.1080/1369183X.1995.9976470

Toma, C. L., & Carlson, C. L. (2015). How do Facebook users believe they come across in their profiles?: A meta-perception approach to investigating Facebook self-presentation. *Communication Research Reports*, *32*(1), 93–101. https://doi.org/10.1080/08824096.2014.990557

Toma, C. L., & Hancock, J. T. (2013). Self-affirmation underlies Facebook use. *Personality and Social Psychology Bulletin*, *39*(3), 321–331. https://doi.org/10.1177/0146167212474694

Tong, S. T., Van Der Heide, B., Langwell, L., & Walther, J. B. (2008). Too much of a good thing? The relationship between number of friends and interpersonal impressions on Facebook. *Journal of Computer-Mediated Communication*, *13*(3), 531–549.

Tsang, A., Wang, K. T., Liu, S., & Yu, L. (2021). Integrating corporate social responsibility criteria into executive compensation and firm innovation: International evidence. *Journal of Corporate Finance*, *70*, 102070.

Turner, J. C., Brown, R. J., & Tajfel, H. (1979). Social comparison and group interest in ingroup favouritism. *European Journal of Social Psychology*, *9*(2), 187–204. https://doi.org/10.1002/ejsp.2420090207

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*(4), 293–315.

United States Census Bureau. (2020). *American housing survey (ahs)* [Retrieved: July 20, 2022]. https://www.census.gov/programs-surveys/ahs/data.html

Utz, S. (2010). Show me your friends and i will tell you what type of person you are: How one's profile, number of friends, and type of friends influence impression formation on social network sites. *Journal of Computer-Mediated Communication*, *15*(2), 314–335. https://doi.org/10.1111/j.1083-6101.2010.01522.x

Utz, S., Tanis, M., & Vermeulen, I. (2012). It's all about being popular: The effects of need for popularity on social network site use. *Cyberpsychology, Behavior, and Social Networking*, *15*(1), 37–42. https://doi.org/10.1089/cyber.2010.0651

van Zoonen, W., & van der Meer, T. (2015). The importance of source and credibility perception in times of crisis: Crisis communication in a socially mediated era. *Journal of Public Relations Research*, *27*(5), 371–388. https://doi.org/10.1080/1062726X.2015.1062382

Walker, D. I. (2022). The economic (in) significance of executive pay esg incentives. *Stan. JL Bus. & Fin.*, *27*, 318.

Walther, J. B., Heide, B. V. D., Hamel, L. M., & Shulman, H. C. (2009). Self-generated versus other-generated statements and impressions in computer-mediated communication: A test of warranting theory using facebook. *Communication Research*, *36*(2), 229–253. https://doi.org/10.1177/0093650208330251

Walther, J. B., Van Der Heide, B., Kim, S.-Y., Westerman, D., & Tong, S. T. (2008). The role of friends' appearance and behavior on evaluations of individuals on facebook: Are we known by the company we keep? *Human Communication Research*, *34*(1), 28–49. https://doi.org/10.1111/j.1468-2958.2007.00312.x

Willis Towers Watson. (2023). *Global report on esg metrics in executive incentive plans* (tech. rep.). Available at https://www.wtwco.com/en-ch/insights/2024/01/global-report-on-esg-metrics-in-incentive-plans-2023.

Wilmot, M. P., & Ones, D. S. (2022). Agreeableness and its consequences: A quantitative review of meta-analytic findings. *Personality and Social Psychology Review*, *26*(3), 242–280. https://doi.org/10.1177/10888683211073007

Wittenbrink, B., & Henly, J. R. (1996). Creating social reality: Informational social influence and the content of stereotypic beliefs. *Personality and Social Psychology Bulletin*, *22*(6), 598–610. https://doi.org/10.1177/0146167296226005

Yarkoni, T. (2010). Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality*, *44*(3), 363–373. https://doi.org/10.1016/j.jrp.2010.04.001

Yu, E. P.-y., Luu, B. V., & Chen, C. H. (2020). Greenwashing in environmental, social and governance disclosures. *Research in International Business and Finance*, *52*, 101192. https://doi.org/10.1016/j.ribaf.2020.101192

Zinovyeva, N., & Bagues, M. (2015). The role of connections in academic promotions. *American Economic Journal: Applied Economics*, *7*(2), 264–92. https://doi.org/10.1257/app.20120337

Zschirnt, E. (2019). Research ethics in correspondence testing: An update. *Research Ethics*, *15*(2), 1–21. https://doi.org/10.1177/1747016118820497

Zschirnt, E., & Ruedin, D. (2016). Ethnic discrimination in hiring decisions: A meta-analysis of correspondence tests 1990–2015. *Journal of Ethnic and Migration Studies*, *42*(7), 1115–1134. https://doi.org/10.1080/1369183X.2015.1133279

Appendix A

Appendix Chapter 2

A.1 Application Text

Hello *advertiser*,

I just saw the ad for the vacant room and I am very interested. My name is *full_name*, I am 24 years old and have recently started my master studies in business administration, which is why I am searching for a room. Because of online lectures, I am still living in Tübingen in a shared apartment, but now I would like to finally move to *city*.

A few words about me: I don't smoke, spent already some time in a shared apartment and know what a cleaning roster is. In my free time I like to meet friends in a cafe or for a cold drink, go jogging, and like to travel. But of course I also like to binge watch my favorite TV series sometimes.

I would be happy to introduce myself and see the room. Time-wise I am open for suggestions.

Kind regards
first_name

P.S.: If you want to see some pictures of me, here is a link to my Instagram profile:
ig_profile_link

Note: The original text was sent in German. Variables are in italics. The last sentence was randomly selected to either be sent with the application text or not.

A.2 Social Networking Site Profiles

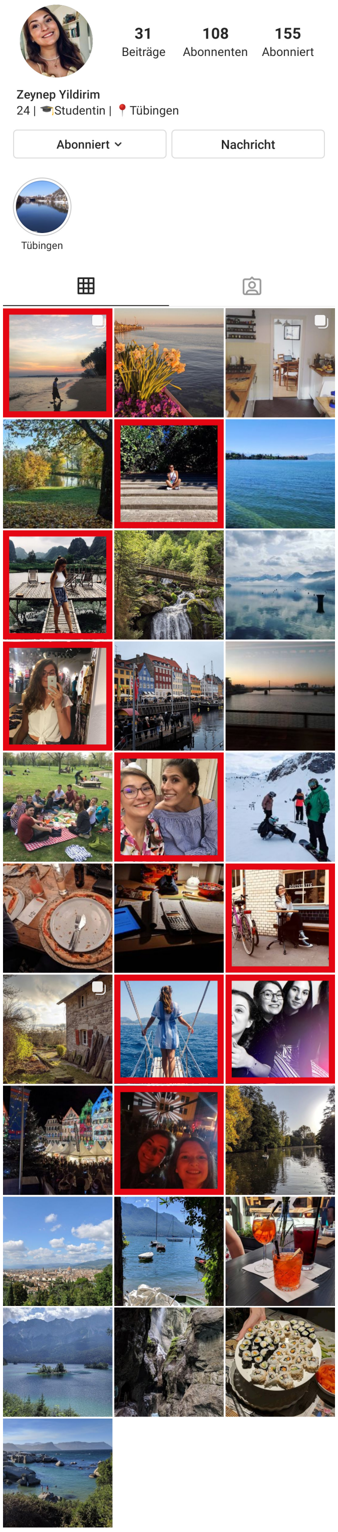


Figure A.1: Example of a Female Profile (non-filler images highlighted in red)

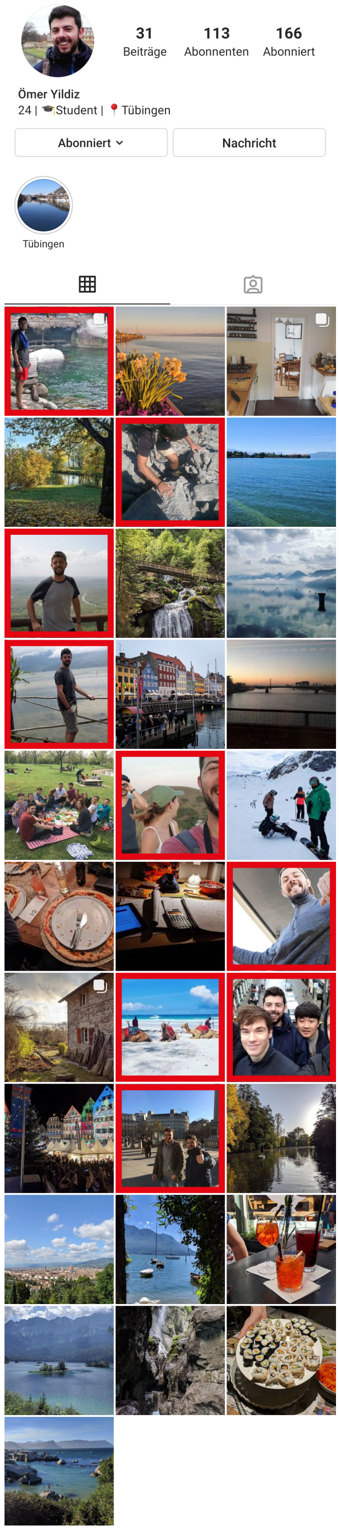


Figure A.2: Example of a Male Profile (non-filler images highlighted in red)

A.3 Tables

Table A.1: Summary Statistics (Selected Variables)

Variable	Mean	SD	Min	Max	N
Turkish	0.495	0.500	0	1	2,810
Female	0.491	0.500	0	1	2,810
Instagram	0.511	0.500	0	1	2,810
Any Response	0.510	0.500	0	1	2,810
Positive Response	0.458	0.498	0	1	2,810
Callback	0.396	0.489	0	1	2,810
Rejection	0.0527	0.223	0	1	2,810
Other Response	0.0616	0.240	0	1	2,810
Call me	0.135	0.342	0	1	2,810
Response Duration (in days)	1.629	4.705	0.00278	127.1	1,434
Response In Same Week	0.950	0.217	0	1	1,434
Room Size (in sqm)	16.94	5.446	4	78	2,810
Total Monthly Rent (in euro)	467.0	157.6	150	1,500	2,810
Number Of Roommates	3.290	1.491	2	17	2,810
Number Of Female Roommates	0.869	0.973	0	7	2,810
Number Of Male Roommates	1.190	1.071	0	9	2,810
Distance To Citycenter (in kilometers)	4.322	3.495	0.100	28.40	2,803
Average Distance To Universities	5.155	3.937	0.400	32.85	2,808
Online Time (in minutes)	118.5	95.44	0	1,020	2,810
Temporary Contract	0.202	0.402	0	1	2,810
Availability (in days)	283.3	302.2	89	3,774	568
Roommates Speak German	0.877	0.328	0	1	2,810
Roommates Speak English	0.717	0.451	0	1	2,810
Roommates Speak Turkish	0.0153	0.123	0	1	2,810
Roommates Speak Arabic	0.0149	0.121	0	1	2,810
Age Range Of Roommates	15.15	22.06	1	98	1,938
Type Of Shared Apartment: Students	0.681	0.466	0	1	2,810
Type Of Shared Apartment: Communally	0.539	0.499	0	1	2,810
Type Of Shared Apartment: Women Only	0.0448	0.207	0	1	2,810
Type Of Shared Apartment: Men Only	0.0178	0.132	0	1	2,810
Type Of Shared Apartment: Employed	0.504	0.500	0	1	2,810
Type Of Shared Apartment: Solitary	0.114	0.318	0	1	2,810
Type Of Shared Apartment: Fraternity	0.0480	0.214	0	1	2,810
Internationals Welcome	0.141	0.348	0	1	2,810
Social Media Requested	0.0915	0.288	0	1	2,810
Female Advertiser	0.453	0.498	0	1	1,349
German Advertiser	0.888	0.315	0	1	1,405
Muslim Advertiser	0.0406	0.197	0	1	1,405
Number Of Smileys In Response	0.431	0.715	0	6	1,434
Viewing Date (duration in days)	2.748	2.889	0.0208	35.79	525

Note: The table shows summary statistics for a selection of variables in the data set.

Table A.2: Summary Statistics – Social Network Site (SNS) Variables

Variable	Mean	SD	Min	Max	N
Postings	31	0	31	31	1,435
Stories	15	0	15	15	1,435
Weekly Followers/Subscribers	109.1	5.317	99	119	1,435
Weekly Subscriptions	169.3	6.720	153	181	1,435
Total Weekly Likes	911.5	62.38	812	1,026	1,435
Total Weekly Likes Of Top3 Images	113.9	20.45	69	149	1,435
Weekly Visits	14.59	9.235	0	43	1,435
Weekly Reach	7.396	4.345	0	18	1,435
Weekly Reach Of Followers/Subscribers	0.521	0.934	0	7	1,435
Weekly Reach Of Non-Subscribers	6.875	4.151	0	17	1,435
Weekly Reach Of Postings	13.05	7.339	0	31	1,435
Weekly Impressions	104.7	75.84	0	304	1,435
Weekly Likes of Postings	2.876	6.093	0	38	1,435

Note: The table shows summary statistics for a selection of SNS variables in the dataset.

Table A.3: Descriptive Results – Responses

	German			Turkish			$\sum$
	with SNS	w/o SNS	$\sum$	with SNS	w/o SNS	$\sum$	
Any Response	394	376	770	356	308	664	1,434
Callback	317	287	604	289	220	509	1,113
Others[1]	40	56	96	34	43	77	173
Call me[2]	102	112	214	101	65	166	380
Rejection	37	33	70	33	45	78	148
Positive Response[3]	357	343	700	323	263	586	1,286
No Response	342	308	650	343	383	726	1,376

Note: The table reports the absolute numbers of responses conditional on the different kinds of responses for ethnicity and gender. [1]"Other" responses include all responses that do not include an explicit invitation to a viewing, rejections excluded. [2]Unlike *Callback*, or any other type of response, *Call me* is non-exclusive. [3]Positive responses exclude rejections.

Table A.4: Results – Probit Models (Average Marginal Effects) – All Ads

	(1)	(2)	(3)	(4)
Callback	Full sample	Full sample	Females	Males
Turkish	-0.114***	-0.120***	-0.101***	-0.146***
	(0.0217)	(0.0164)	(0.0319)	(0.0205)
Female	0.120***	0.113***	-	-
	(0.0108)	(0.0115)		
Instagram	0.0194	0.0216	0.0376	0.00569
	(0.0294)	(0.0254)	(0.0435)	(0.0251)
Turkish × Instagram	0.0676*	0.0809***	0.0616	0.108***
	(0.0382)	(0.0306)	(0.0571)	(0.0328)
Observations	3,676	3,335	1,776	1,559
City Control Variables	No	Yes	Yes	Yes
Room and Apartment Control Variables	No	Yes	Yes	Yes
Covid-19 Control Variables	No	Yes	Yes	Yes
District Demographic Control Variables	No	Yes	Yes	Yes
Time Control Variables	No	Yes	Yes	Yes
Pseudo R^2	0.0193	0.135	0.126	0.156

Note: The table reports the average marginal effects of probit regression models with a callback dummy as the dependent variable. The sample includes all vacant room ads, also those that specify the gender of the roommate that is looked for. Column 1 reports the main effects without control variables, column 2 shows the main effects including controls. Columns 3 and 4 report the main effects including control variables, split by gender. Testing for differences between the coefficients of *Turkish*, *Instagram*, and *Turkish* × *Instagram* across the two subsamples in column (3) and (4), we cannot reject the hypothesis that the coefficients are equal. Standard errors (in parentheses) are clustered at the city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Results – Probit Models (Average Marginal Effects) – Either Females Or Males Accepted

Callback	(1) Full sample	(2) Full sample	(3) Females	(4) Males
Turkish	-0.135***	-0.144***	-0.121*	-0.254***
	(0.0397)	(0.0391)	(0.0631)	(0.0603)
Female	0.160***	0.130***	-	-
	(0.0418)	(0.0413)		
Instagram	0.0621	0.0617	0.0802	0.0226
	(0.0669)	(0.0541)	(0.0663)	(0.0642)
Turkish × Instagram	-0.0208	-0.00381	0.0152	-0.0316
	(0.0885)	(0.0759)	(0.104)	(0.0990)
Observations	866	785	522	250
City Control Variables	No	Yes	Yes	Yes
Room and Apartment Control Variables	No	Yes	Yes	Yes
Covid-19 Control Variables	No	Yes	Yes	Yes
District Demographic Control Variables	No	Yes	Yes	Yes
Time Control Variables	No	Yes	Yes	Yes
Pseudo R^2	0.0368	0.147	0.157	0.260

Note: The table reports the average marginal effects of probit regression models with a callback dummy as the dependent variable. The sample consists of vacant room ads that only accept one gender. Column 1 reports the main effects without control variables, column 2 shows the main effects including controls. Columns 3 and 4 report the main effects including control variables, split by gender. Testing for differences between the coefficients of *Turkish* across the two subsamples in column (3) and (4), we can reject the hypothesis that the coefficients are equal (p = 0.0096), which is not the case for *Instagram* and *Turkish* × *Instagram*. Standard errors (in parentheses) are clustered at the city level. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A.6: Results – Probit Models (Average Marginal Effects) – Positive Response

Positive Response	(1) Full sample	(2) Full sample	(3) Females	(4) Males
Turkish	-0.123***	-0.124***	-0.111***	-0.145***
	(0.0234)	(0.0187)	(0.0297)	(0.0264)
Female	0.104***	0.0908***		
	(0.0118)	(0.0130)	-	-
Instagram	-0.0166	-0.00977	-0.0147	-0.0102
	(0.0231)	(0.0211)	(0.0348)	(0.0291)
Turkish × Instagram	0.100***	0.105***	0.0979**	0.122***
	(0.0347)	(0.0260)	(0.0440)	(0.0422)
Observations	2,810	2,550	1,254	1,296
City Control Variables	No	Yes	Yes	Yes
Room and Apartment Control Variables	No	Yes	Yes	Yes
Covid-19 Control Variables	No	Yes	Yes	Yes
District Demographic Control Variables	No	Yes	Yes	Yes
Time Control Variables	No	Yes	Yes	Yes
Pseudo R^2	0.0144	0.149	0.155	0.159

Note: The table reports the average marginal effects of probit regression models with a positive response dummy as the dependent variable. A positive response includes both callbacks and *other* responses, i.e., advertisers requesting additional information. The sample consists of vacant room ads that do not impose any restriction on an applicant's gender. Column 1 reports the main effects without control variables, column 2 shows the main effects including controls. Columns 3 and 4 report the main effects including control variables, split by gender. Testing for differences between the coefficients of *Turkish*, *Instagram*, and *Turkish* × *Instagram* across the two subsamples in column (3) and (4), we cannot reject the hypothesis that the coefficients are equal. Standard errors (in parentheses) are clustered at the city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: Results – Probit Models (Average Marginal Effects) – Rejection

Rejection	(1) Full sample	(2) Full sample	(3) Females	(4) Males
Turkish	0.0160	0.0197	0.0108	0.0376**
	(0.0131)	(0.0137)	(0.0176)	(0.0180)
Female	-0.00826	-0.00760	-	-
	(0.0104)	(0.0118)		
Instagram	0.00195	0.00246	-0.0155	0.0219
	(0.0120)	(0.0121)	(0.0158)	(0.0144)
Turkish × Instagram	-0.0191	-0.0261	-0.0124	-0.0447*
	(0.0198)	(0.0191)	(0.0213)	(0.0259)
Observations	2,810	2,550	1,254	1,240
City Control Variables	No	Yes	Yes	Yes
Room and Apartment Control Variables	No	Yes	Yes	Yes
Covid-19 Control Variables	No	Yes	Yes	Yes
District Demographic Control Variables	No	Yes	Yes	Yes
Time Control Variables	No	Yes	Yes	Yes
Pseudo R^2	0.00324	0.0831	0.117	0.148

Note: The table reports the average marginal effects of probit regression models with a rejection dummy as the dependent variable. The sample consists of vacant room ads that do not impose any restriction on an applicant's gender. Column 1 reports the main effects without control variables, column 2 shows the main effects including controls. Columns 3 and 4 report the main effects including control variables, split by gender. Testing for differences between the coefficients of *Instagram* across the two subsamples in column (3) and (4), we can reject the hypothesis that the coefficients are equal ($p = 0.0522$), which is not the case for *Turkish* and *Turkish* × *Instagram*. Standard errors (in parentheses) are clustered at the city level. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A.8: Results – Probit Models (Average Marginal Effects) – Any Response

Any Response	(1) Full sample	(2) Full sample	(3) Females	(4) Males
Turkish	-0.105***	-0.103***	-0.103***	-0.109***
	(0.0223)	(0.0230)	(0.0310)	(0.0420)
Female	0.0962***	0.0837***	-	-
	(0.0111)	(0.0101)		
Instagram	-0.0148	-0.0106	-0.0308	0.00650
	(0.0239)	(0.0227)	(0.0419)	(0.0306)
Turkish × Instagram	0.0792*	0.0809**	0.0877	0.0805
	(0.0421)	(0.0386)	(0.0568)	(0.0625)
Observations	2,810	2,550	1,254	1,296
City Control Variables	No	Yes	Yes	Yes
Room and Apartment Control Variables	No	Yes	Yes	Yes
Covid-19 Control Variables	No	Yes	Yes	Yes
District Demographic Control Variables	No	Yes	Yes	Yes
Time Control Variables	No	Yes	Yes	Yes
Pseudo R^2	0.0113	0.136	0.154	0.138

Note: The table reports the average marginal effects of probit regression models with an any response dummy (including callbacks, others, rejections) as the dependent variable. If the dependent variable *any_response* equals zero, the respective application did not receive a callback. The sample consists of vacant room ads that do not impose any restriction on an applicant's gender. Column 1 reports the main effects without control variables, column 2 shows the main effects including controls. Columns 3 and 4 report the main effects including control variables, split by gender. Testing for differences between the coefficients of *Turkish*, *Instagram*, and *Turkish* × *Instagram* across the two subsamples in column (3) and (4), we cannot reject the hypothesis that the coefficients are equal. Standard errors (in parentheses) are clustered at the city level. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A.9: Results – Probit Models (Average Marginal Effects) – Three-Way Interaction

Callback	(1)	(2)	(3)	(4)
Turkish	-0.0594***	-0.0545***	-0.127***	-0.119***
	(0.0210)	(0.0191)	(0.0391)	(0.0252)
Female	0.107***	0.0964***	0.101***	0.0804***
	(0.00964)	(0.0116)	(0.0298)	(0.0288)
Instagram	0.0533***	0.0641***	0.00834	0.0136
	(0.0183)	(0.0166)	(0.0462)	(0.0362)
Turkish × Female	-	-	0.0409	0.0419
			(0.0493)	(0.0430)
Turkish × Instagram	-	-	0.120*	0.114**
			(0.0711)	(0.0469)
Female × Instagram	-	-	0.00423	0.0233
			(0.0630)	(0.0454)
Turkish × Female × Instagram	-	-	-0.0611	-0.0657
			(0.0813)	(0.0639)
Observations	2,810	2,550	2,810	2,550
City Control Variables	No	Yes	No	Yes
Room and Apartment Control Variables	No	Yes	No	Yes
Covid-19 Control Variables	No	Yes	No	Yes
District Demographic Control Variables	No	Yes	No	Yes
Time Control Variables	No	Yes	No	Yes
Pseudo R^2	0.0140	0.183	0.0159	0.185

Note: The table reports the average marginal effects of probit regression models with a callback dummy as the dependent variable. The sample consists of vacant room ads that do not impose any restriction on an applicant's gender. Column 1 reports the main effects without control variables and interaction effects. Column 2 shows the main effects including control variables, column 3 shows the main and interaction effects without control variables, and column 4 shows main and interaction effects including controls. Standard errors (in parentheses) are clustered at the city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.10: Results – Probit Models (Average Marginal Effects) – SNS Variables

Callback	(1) SNS subsample	(2) SNS Turkish subsample	(3) SNS German subsample	(4) SNS Female subsample	(5) SNS Male subsample
Turkish	-0.0726**	-	-	-0.0834	-0.0963
	(0.0348)			(0.0899)	(0.0657)
Female	0.0580**	0.0862**	-0.00348	-	-
	(0.0289)	(0.0368)	(0.0995)		
Weekly Followers/Subscribers	0.00928***	0.0168*	0.00953***	0.0124	0.00942*
	(0.00344)	(0.00886)	(0.00340)	(0.00851)	(0.00545)
Weekly Subscriptions	-0.00506	-0.00236	-0.00971	-0.00727	-0.00584
	(0.00320)	(0.00396)	(0.00689)	(0.00578)	(0.00560)
Total Weekly Likes	-0.000262	-0.00108	-2.33e-05	-0.000127	-0.000375
	(0.000262)	(0.000761)	(0.000509)	(0.000445)	(0.000477)
Weekly Reach	-0.0474***	-0.0390**	-0.0747***	-0.0408	-0.0724**
	(0.0133)	(0.0191)	(0.0236)	(0.0274)	(0.0292)
Weekly Reach Of Non-Subscribers	0.0330***	0.0414**	0.0459**	0.0395***	0.0207
	(0.00743)	(0.0205)	(0.0192)	(0.0140)	(0.0264)
Weekly Reach Of Postings	0.0114	0.000315	0.0232**	0.00371	0.0364***
	(0.00909)	(0.00818)	(0.0115)	(0.0135)	(0.0116)
Weekly Impressions	0.000253	0.000232	0.000167	0.000425	-0.000561
	(0.000490)	(0.000536)	(0.000796)	(0.000613)	(0.000907)
Weekly Likes of Postings	-0.00342	-0.00848*	-0.00147	-0.00194	-0.00986**
	(0.00251)	(0.00447)	(0.00354)	(0.00331)	(0.00503)
Number of Applications	-0.000795	-0.00589	0.00351	-0.00328	0.000456
	(0.00226)	(0.00377)	(0.00295)	(0.00390)	(0.00299)
Number of SNS Applications	-0.00108	0.00787	-0.00918*	0.00161	-0.00129
	(0.00515)	(0.00746)	(0.00558)	(0.00663)	(0.00713)
Observations	1,435	699	736	703	732
City Control Variables	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.0844	0.0950	0.0969	0.0921	0.0762

Note: The table shows the average marginal effects of probit regression models with a callback dummy as the dependent variable. The sample consists of applications that include an SNS profile link for vacant room ads that do not impose a restriction on an applicant's gender. Column 1 shows the SNS effects for the entire SNS sample. Column 2 shows the main effects for a subsample consisting of SNS applications using Turkish-sounding names, while column 3 reports the main effects for German-sounding names. Columns 4-5 show the effects for subsamples of female and male names, respectively. Number of applications and number of SNS applications are reported on name and week level (for a given name in a given week). Standard errors (in parentheses) are clustered at the city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.11: Results – OLS Models – SNS Visit Variables

	(1) Weekly Visits	(2) Weekly Impressions	(3) Weekly Reach of Non-Subscribers
Number Of SNS Applications	0.470***	4.792***	0.272***
	(0.0353)	(0.291)	(0.0143)
Turkish	0.295	2.755	0.380
	(0.576)	(5.009)	(0.333)
Female	11.94***	71.22***	3.846***
	(0.587)	(5.871)	(0.282)
Turkish × Female	-1.491	-9.320*	-1.720***
	(0.896)	(5.240)	(0.297)
Weekly Followers	0.205***	2.055***	0.115***
	(0.0518)	(0.561)	(0.0383)
Weekly Subscriptions	0.359***	2.030***	0.156***
	(0.0488)	(0.326)	(0.0239)
Constant	-83.31***	-582.0***	-38.15***
	(5.992)	(27.59)	(1.724)
Observations	1,435	1,435	1,435
City Control Variables	Yes	Yes	Yes
R-squared	0.587	0.472	0.451

Note: The table reports the results of different OLS models using weekly visits, weekly impressions, and weekly reach of non-subscribers as dependent variables. The sample consists of vacant room ads that do not impose any restriction on an applicant's gender. Number of SNS applications reported on name and week level (for a given name in a given week). An SNS application contains a link to an SNS profile. Standard errors (in parentheses) are clustered at the city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.12: Results – OLS Models – Others

	(1) Number of Smileys In Response	(2) Sentiment Of Response
Turkish	0.0102	-0.00753
	(0.0471)	(0.0102)
Female	0.110**	0.0296***
	(0.0436)	(0.00955)
Instagram	0.0994**	0.0263*
	(0.0400)	(0.0137)
Female Advertiser	0.168***	0.0102
	(0.0346)	(0.0123)
German Advertiser	0.0942	0.0719***
	(0.0583)	(0.0230)
Constant	0.268***	0.419***
	(0.0722)	(0.0245)
Observations	1,345	1,344
City Control Variables	Yes	Yes
R-squared	0.066	0.052

Note: The table reports the results of two OLS models using the number of smileys in the response and the sentiment of the response as dependent variables. The sample consists of vacant room ads that do not impose any restriction on an applicant's gender. Sentiment data is collected using the Google Cloud Natural Language API. The score ranges from -1.0 (negative) to 1.0 (positive) and reflects the emotional opinion of the text being negative, neutral, or positive, see: https://cloud.google.com/natural-language/ [Retrieved: August 2, 2022]. Standard errors (in parentheses) are clustered at the city level. * p < 0.1, ** p < 0.05, *** p < 0.01.

A.4 Figures

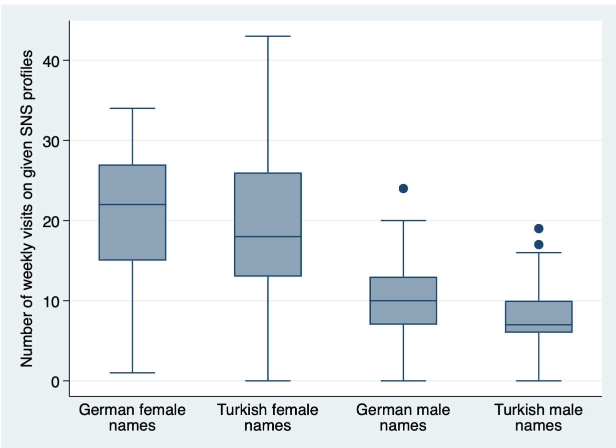


Figure A.3: Weekly SNS Profile Visits by Ethnicity and Gender During the Experiment

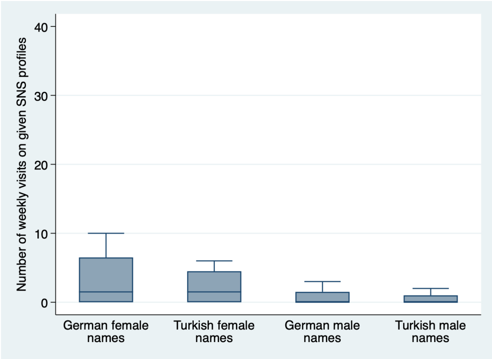


Figure A.4: Weekly SNS Profile Visits by Ethnicity and Gender Before the Start of the Experiment

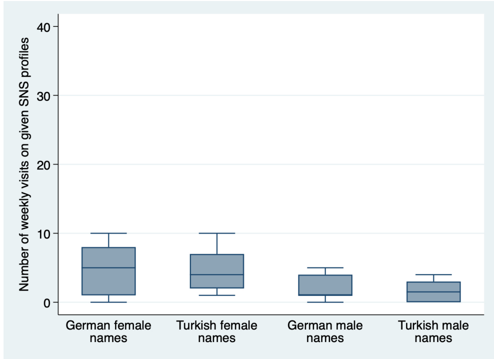


Figure A.5: Weekly SNS Profile Visits by Ethnicity and Gender After the End of the Experiment

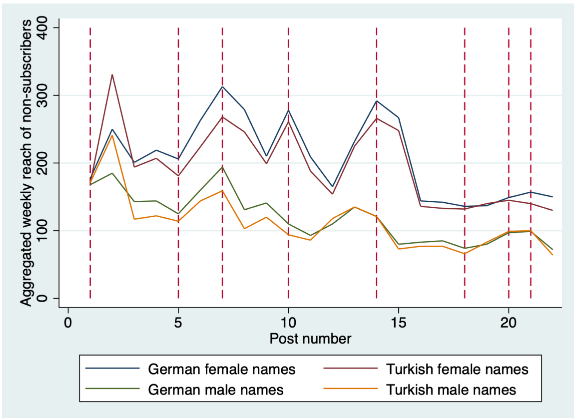


Figure A.6: Aggregated Weekly Reach of Non-Subscribers per Posting by Ethnicity and Gender
(dashed lines indicate the non-filler postings)

A.5 Online Pilot Experiments

Prior to conducting the main experiment, we run a randomized online pilot experiment in order to assess the perceived origin and other characteristics of potential male and female fictitious profile owners. Additionally, we conduct a separate survey to enhance and evaluate the profiles' perceived realism. Following these, we carry out another randomized online pilot to verify the efficacy of the treatment conditions in manipulating relevant beliefs, ensuring that the selected names clearly and unambiguously represent the intended gender and ethnicity in the absence of additional social media information. This final pilot was conducted after the main experiment, following an anonymous reviewer's recommendation.

A.5.1 First Online Pilot Experiment

To create realistic and robust social media profiles, we recruit five students from our university (two female, three male) who meet the criteria of being either of Turkish origin, second-generation immigrants from Turkey, or perceived as having Turkish origin. In addition, the students need to be perceived as both Turkish *and* German, since they are randomly assigned to either a Turkish or German name. Because we intend to compare outcomes for both female and male profile owners, we aim for two individuals who were as similar as possible in terms of perceived ethnicity, personality traits, and suitability as potential roommates, thereby minimizing potential confounding factors.

Accordingly, the primary objective of the first online pilot experiment is to evaluate perceived origin and other characteristics of potential male and female fictitious profile owners. This assessment enables us to select one female-male pair closely matched on all attributes, particularly ensuring that both are perceived as having a mixed Turkish and German background.

We conduct the experiment in November 2019 with $n = 158$ students. Each subject is randomly assigned to receive a random selection of images from three of the five potential fictitious profile owners. To ensure authenticity and relevance, the "models" provided personal photos that they had previously shared on social media, guaranteeing that the images are suitable for this experiment. These same photos, among others, are later used in constructing the experimental profiles.

Survey Design After a short introductory text explaining the study's focus on perceptions of social media images – deliberately avoiding any mention of race, ethnicity, or origin – we randomly assign participants to view a subset of images. Each participant is given images from one of the two female and two of the three male candidates, presented in random order.

For each candidate, participants are asked to carefully examine the respective images and then rate the individuals' personality in terms of conscientiousness, agreeableness, and extraversion, using the Big Five Inventory-2-XS (Rammstedt et al., 2020; Soto & John, 2017). This inventory includes three items per trait, rated on a 7-point Likert scale (ranging from "fully applies" to "does not apply at all").[1]

Subsequently, participants are asked to rate the likelihood that each individual originated from specific countries or regions, including Turkey and Germany, together with additional control items (Poland, Italy, and North Africa), using a five-point Likert scale (ranging from "unlikely" to "likely"). Afterwards, we ask subjects whether they could envision sharing an apartment with the individual as a potential roommate, using another five-point Likert scale ranging from "yes, definitely" to "no, definitely not" (see Figure A.7).

Finally, participants are asked to provide demographic information, including gender, age, education, household size, migration background, and their frequency of using various social media platforms.

Results Table A.13 provides an overview of the participants' demographic characteristics. Approximately 61% of the respondents identify as female, and about 56% are between 21 and 29 years old. Half of the sample report living in a shared apartment, with an average household size of three. Most participants hold either a university degree (presumably master's students) or a high school diploma (bachelor's students). Approximately 18% of respondents report having a migration background. In terms of social media usage, Instagram and Youtube are the platforms used most frequently.

Regarding the potential profile owners' characteristics, the main purpose of this online experiment, Table A.14 reveals notable differences in perceived Turkish vs. German origin. Among the female candidates, *Female_2* (second generation Turkish immigrant) shows a relatively balanced perception (3.1 vs. 3.6), whereas *Female_1* (Turkish native) displays a large discrepancy (3.9 vs. 2.9). For the male candidates, *Male_2* (second generation Turkish immigrant) is the most balanced (3.6 vs. 3.5), followed by *Male_3* (German native; 3.0 vs. 3.7). In contrast, *Male_1* (second generation Turkish immigrant) exhibits the greatest gap (2.6 vs. 4.2).

Although *Male_2* reveals the best fit in terms of balanced perceived origin, *Male_3* aligns more closely with *Female_2* on both dimensions of perceived Turkish and German backgrounds,

[1] The Big Five Inventory (Soto & John, 2017) measures five core personality dimensions – openness, conscientiousness, extraversion, agreeableness, and neuroticism. The BFI-2-XS is a very short version of the BFI introduced by Rammstedt et al. (2020). It demonstrates high reliability and validity.

as evidenced by their nearly identical scores (3.1 vs. 3.0 for Turkish and 3.6 vs. 3.7 for German). Additionally, their perceived personalities in terms of agreeableness and extraversion, as well as suitability as potential roommates are closely matched.

We therefore select *Female_2* and *Male_3* as the fictitious profile owners. Both are highly homogeneous across all characteristics, ensuring negligible differences apart from gender. Most importantly, both are perceived to be of mixed Turkish-German origin, which makes our design – focusing on ethnicity as the sole differentiating factor between profiles – particularly effective.

20% ausgefüllt

Betrachten Sie bitte folgende Bilder der gleichen Person:

1. Wie würden Sie die Person auf den Bildern einschätzen?

Die Person auf dem Bild...	Trifft überhaupt nicht zu						Trifft voll zu
...arbeitet gründlich	○	○	○	○	○	○	○
...ist kommunikativ und gesprächig	○	○	○	○	○	○	○
...ist manchmal etwas grob zu anderen	○	○	○	○	○	○	○
...kann verzeihen	○	○	○	○	○	○	○
...ist eher faul	○	○	○	○	○	○	○
...kann aus sich herausgehen, ist gesellig	○	○	○	○	○	○	○
...erledigt Aufgaben wirksam und effizient	○	○	○	○	○	○	○
...ist zurückhaltend	○	○	○	○	○	○	○
...geht rücksichtsvoll und freundlich mit anderen um	○	○	○	○	○	○	○

2. Für wie wahrscheinlich halten Sie es, dass die Vorfahren der abgebildeten Person aus einem der folgenden Länder/Regionen stammen? (Mehrfachauswahl möglich)

	Unwahrscheinlich				Wahrscheinlich
Deutschland	○	○	○	○	○
Polen	○	○	○	○	○
Türkei	○	○	○	○	○
Italien	○	○	○	○	○
Nordafrika	○	○	○	○	○

3. Auch wenn Ihnen nur Bilder der Person zur Verfügung stehen: Könnten Sie sich grundsätzlich vorstellen, mit der Person gemeinsam in einer Wohngemeinschaft (WG) zu leben?

○ Ja, auf jeden Fall
○ Eher ja
○ Unentschlossen
○ Eher nein
○ Nein, auf keinen Fall

Zurück Weiter

Raphael Moritz, Eberhard Karls Universität Tübingen

Figure A.7: Screenshot of the Evaluation Task (*Female_1*)

Table A.13: First Online Pilot Experiment: Demographics (Summary Statistics)

	Mean	S.D.
Female	0.609	0.489
Age <17	0.013	0.114
Age 18-20	0.170	0.376
Age 21-29	0.559	0.497
Age 30-39	0.135	0.342
Age 40-49	0.020	0.139
Age 50-59	0.052	0.223
Age >60	0.052	0.223
Living in Shared Apartment	0.484	0.500
Household Members	3.002	1.993
No Education	0.000	0.000
Education: Primary School	0.007	0.081
Education: Intermediate School	0.013	0.114
Education: Gymnasium	0.402	0.491
Education: Vocational Training	0.065	0.247
Education: University College Degree	0.065	0.247
Education: University Degree	0.422	0.494
Education: PhD	0.020	0.139
Migration Background	0.184	0.388
Social Media Usage: Facebook	2.952	1.484
Social Media Usage: Instagram	3.431	1.799
Social Media Usage: Twitter	1.551	1.140
Social Media Usage: Snapchat	1.871	1.454
Social Media Usage: Pinterest	1.442	0.857
Social Media Usage: Xing	1.457	0.834
Social Media Usage: Linkedin	1.676	1.051
Social Media Usage: Youtube	3.785	1.035
Obs.	158	

Note: The table reports summary statistics of the demographic characteristics of participants of the first online pilot experiment. All variables are dummies, except for the number of household members and variables indicating social media usage for given platforms, which are measured on a 5-point Likert scale, with 1 indicating never and 5 indicating several times a day.

Table A.14: First Online Pilot Experiment: Summary Statistics on Potential Profile Owner's Characteristics

	(1)		(2)		(3)		(4)		(5)	
	Female_1		*Female_2*		*Male_1*		*Male_2*		*Male_3*	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Perceived Origin										
German	2.894	1.140	3.606	1.202	4.165	0.924	3.459	1.181	3.717	0.987
Polish	1.798	0.979	2.596	1.212	2.365	1.223	2.247	1.204	2.609	1.099
Turkish	3.904	0.876	3.117	1.217	2.647	1.279	3.553	1.041	3.022	1.204
Italian	3.904	0.971	3.521	1.133	3.412	1.083	3.529	0.921	3.533	1.143
North Africa	2.452	1.299	1.851	1.077	1.576	1.051	2.282	1.287	1.728	1.017
Personality Traits										
Conscientiousness	4.981	1.113	4.901	0.934	5.145	0.997	5.090	1.016	4.464	0.914
Agreeableness	5.554	0.881	5.067	0.957	4.682	1.000	5.176	0.860	5.159	0.929
Extraversion	5.801	0.966	5.660	0.904	5.314	0.960	4.518	1.218	5.496	1.221
Perceived Fit										
Shared Apartment Fit	4.029	1.083	3.798	0.934	3.306	1.091	3.565	1.005	3.685	1.037
Obs.	104		94		85		85		92	

Note: The items pertaining to perceived origin are measured on a 5-point Likert scale, with 1 indicating *unlikely* and 5 indicating *likely*. The personality traits are measured on a 7-point Likert scale, with responses ranging from *does not apply at all* (1) to *fully applies* (7). The perceived fit as a potential roommate is measured on a 5-point Likert scale, with 1 indicating *no, under no circumstances* and 5 indicating *yes, definitely.*

A.5.2 Second Online Pilot Experiment

The aim of the second online pilot experiment is to validate the attribution of the eight treatment names to the intended ethnic origin and gender. The experiment was conducted online using a professional survey platform in June 2023. We recruit a total of $n = 1,725$ participants via an university-wide circular mail.[2]

Survey Design After a brief introduction that deliberately avoided mentioning ethnicity, origin, or related topics, participants are randomly assigned two of the eight names of our experiment (between-subjects design). Subjects are first asked to indicate the perceived origin of the randomly selected names individually using open-ended questions. Subsequently, subjects are requested to rate both names on gender, origin, and religious affiliation using predefined items on a 7-point Likert scale ranging from very unlikely to very likely (see Figure A.8). Finally, participants provide demographic information, including gender, age, occupation, household size, experience living in shared apartments, nationality, and whether they had a migration background – specifically, if one or both parents emigrated to Germany.

Results We exclude 324 observations from the analysis as participants either complete the questionnaire too quickly or report not having answered truthfully. In terms of demographic characteristics, the participants of this online pilot experiment closely resemble the sample of the field experiment. Table A.15 shows that about 67.4% are students, and the average age is 28.7 years (median age is 24 years). Approximately 73.5% identify as female. Most participants (66.4%) have experience living in a shared apartment and have German nationality, while 25.8% report having a migration background.

The results, presented in Table A.16, indicate that the randomized treatment effectively manipulates participants' beliefs, with Turkish names predominantly perceived as Turkish in both the open-ended responses and the 7-point Likert-scale ratings. A similar pattern emerges for German names (see columns 1 and 4 of Table A.16). Although Turkish names may also evoke

[2]Note that the first stage of this experiment serves as a manipulation check for Chapter 2, while a second and third stage – during which participants are asked about commonly held stereotypes – are also conducted for Chapter 3 (see Section B.6.1, p. 219).

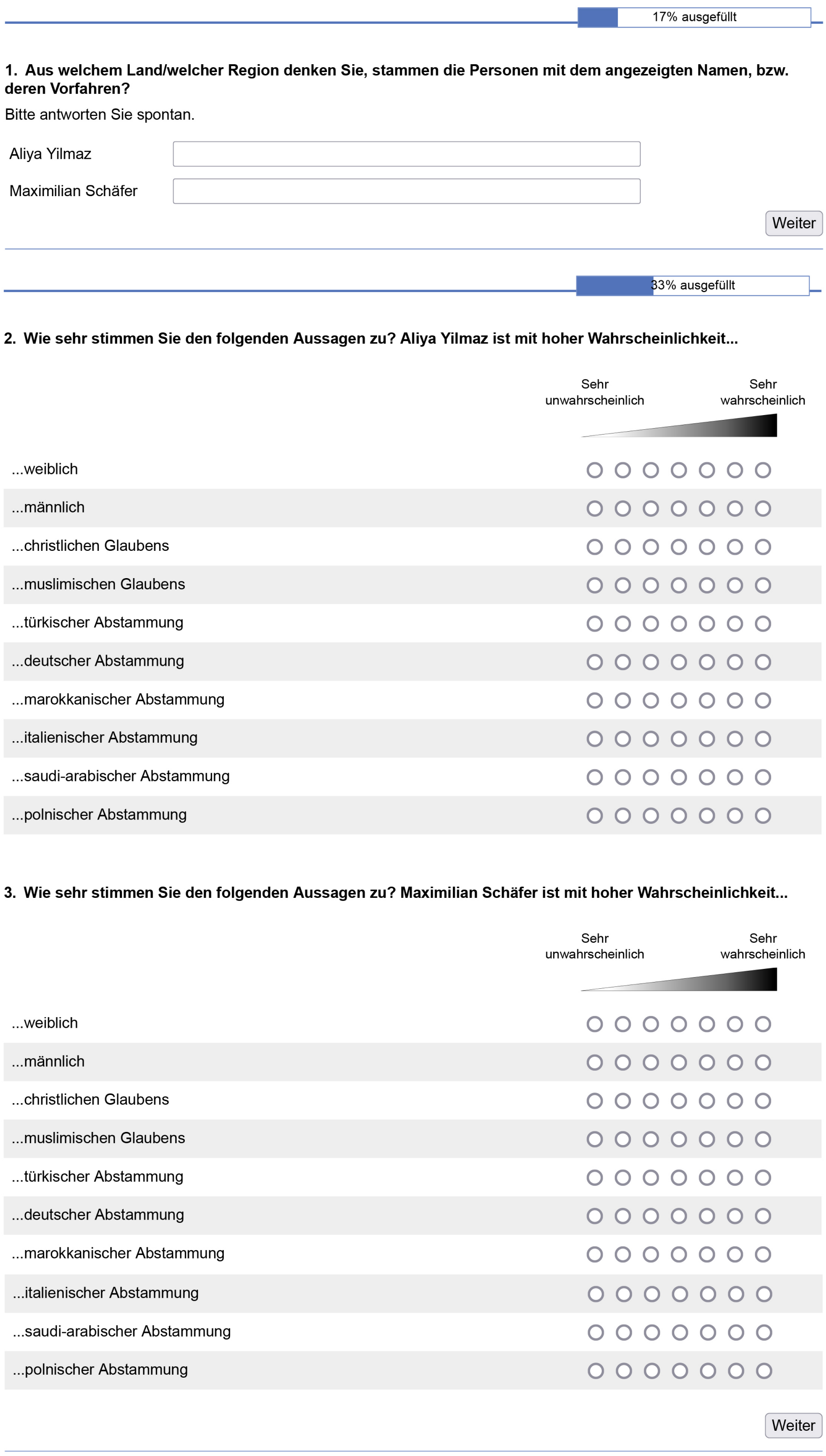
associations with Arabic or Middle Eastern origins, Turkish consistently exhibits the highest mean values across all conditions.

Moreover, female (male) names are significantly more likely to be perceived as female (male), and Turkish (German) names are more frequently associated with Muslim (Christian) religious affiliations. Two-sample t-tests confirm that the mean differences in characteristics between Turkish and German names are statistically significant.

Table A.17 further supports these findings by examining the proportions of participants who perceived each origin or gender as very (un)likely or quite (un)likely. Female (male) names were correctly classified as female (male) in the vast majority of cases (87.4%–99.1%), and Turkish-sounding (German-sounding) names were correctly categorized with respect to ethnicity nearly as frequently (93.1%–99.7%).

These results are confirmed by OLS regressions of the manually coded open-ended responses about the perceived origin of the treatment name on a dummy variable indicates whether the participant was randomly assigned a Turkish name, along with selected interaction terms (see Table A.18). The findings demonstrate that participants strongly associate the Turkish name with a Turkish origin, as intended ($\beta = 0.808$, $p = 0.000$; see column 1 of Table A.18). In contrast, the treatment effects on other potential origins, such as Arabic or Middle Eastern (columns 2 and 3), are substantially lower. All models control for demographic characteristics (e.g., age, occupation, nationality/migration background) as well as additional survey controls.

Overall, it is highly unlikely that participants in the field experiment would fail to associate the intended ethnicity and/or gender.



17% ausgefüllt

1. Aus welchem Land/welcher Region denken Sie, stammen die Personen mit dem angezeigten Namen, bzw. deren Vorfahren?

Bitte antworten Sie spontan.

Aliya Yilmaz

Maximilian Schäfer

Weiter

33% ausgefüllt

2. Wie sehr stimmen Sie den folgenden Aussagen zu? Aliya Yilmaz ist mit hoher Wahrscheinlichkeit...

	Sehr unwahrscheinlich						Sehr wahrscheinlich
...weiblich	○	○	○	○	○	○	○
...männlich	○	○	○	○	○	○	○
...christlichen Glaubens	○	○	○	○	○	○	○
...muslimischen Glaubens	○	○	○	○	○	○	○
...türkischer Abstammung	○	○	○	○	○	○	○
...deutscher Abstammung	○	○	○	○	○	○	○
...marokkanischer Abstammung	○	○	○	○	○	○	○
...italienischer Abstammung	○	○	○	○	○	○	○
...saudi-arabischer Abstammung	○	○	○	○	○	○	○
...polnischer Abstammung	○	○	○	○	○	○	○

3. Wie sehr stimmen Sie den folgenden Aussagen zu? Maximilian Schäfer ist mit hoher Wahrscheinlichkeit...

	Sehr unwahrscheinlich						Sehr wahrscheinlich
...weiblich	○	○	○	○	○	○	○
...männlich	○	○	○	○	○	○	○
...christlichen Glaubens	○	○	○	○	○	○	○
...muslimischen Glaubens	○	○	○	○	○	○	○
...türkischer Abstammung	○	○	○	○	○	○	○
...deutscher Abstammung	○	○	○	○	○	○	○
...marokkanischer Abstammung	○	○	○	○	○	○	○
...italienischer Abstammung	○	○	○	○	○	○	○
...saudi-arabischer Abstammung	○	○	○	○	○	○	○
...polnischer Abstammung	○	○	○	○	○	○	○

Weiter

Figure A.8: Screenshot of the Evaluation Task (Second Online Pilot Experiment)

Table A.15: Second Online Pilot Experiment: Summary Statistics

	Mean	S.D.	Min	Max	Obs.
Treatment Variables					
Treatment 1: Turkish name first	0.481	0.499	0	1	1,401
Treatment 2: Turkish ethnicity	0.498	0.500	0	1	1,401
Treatment 3: Turkish name	0.502	0.500	0	1	1,401
Participant's Demographics					
Female	0.735	0.441	0	1	1,262
Age	28.77	11.57	16	79	1,200
Student	0.674	0.469	0	1	1,265
Employed (full-time)	0.169	0.375	0	1	1,265
Employed (part-time)	0.117	0.322	0	1	1,265
Shared apartment experience	0.664	0.473	0	1	1,255
Shared apartment experience (in years)	3.833	3.160	1	25	797
Number of roommates	3.759	4.638	1	53	1,214
German	0.924	0.265	0	1	1,201
Turkish	0.0236	0.152	0	1	1,401
Migration background	0.258	0.438	0	1	1,223
Survey Variables					
Time (in seconds)	423.5	171.8	105	1,168	1,401
Finished	0.903	0.296	0	1	1,401
Share of missing responses	2.525	2.559	0	35	1,401

Note: The table reports summary statistics for the first online pilot experiment.

Table A.16: Second Online Pilot Experiment: Summary Statistics (Name's Characteristics)

	(1) Mean	(2) S.D.	(3) Obs.	(4) Mean	(5) S.D.	(6) Obs.	(7) Diff. in Means (p-value)
	Turkish Names			German Names			
Open-ended Question							
German origin	0.0279	0.165	1,397	0.965	0.184	1,401	-0.94*** (0.000)
Turkish origin	0.812	0.391	1,397	0.00286	0.0534	1,401	0.81*** (0.000)
Arabic origin	0.0852	0.279	1,397	0.00143	0.0378	1,401	0.84*** (0.000)
Middle Eastern origin	0.0379	0.191	1,397	0	0	1,401	0.04*** (0.000)
Other origin	0.0372	0.189	1,397	0.0307	0.173	1,401	0.01 (0.340)
7-point Likert Scale							
Turkish	5.891	1.082	1,401	1.930	1.063	1,401	3.96*** (0.000)
German	2.924	1.414	1,401	6.174	1.027	1,401	-3.25*** (0.000)
Saudi Arabian	3.774	1.731	1,400	1.771	0.958	1,401	2.00*** (0.000)
Moroccan	3.545	1.624	1,401	1.858	0.982	1,401	1.69*** (0.000)
Italian	2.040	1.032	1,401	2.393	1.266	1,401	-0.35*** (0.000)
Polish	1.887	1.032	1,401	2.937	1.427	1,401	-1.05*** (0.000)
Christian	2.665	1.138	1,401	4.829	1.101	1,401	-2.16*** (0.000)
Muslim	5.546	1.069	1,401	2.273	1.105	1,401	3.27*** (0.000)
	Female Names			Male Names			
Female	6.318	1.380	1,404	1.349	0.848	1,398	4.97*** (0.000)
Male	1.747	1.392	1,401	6.732	0.754	1,398	-4.99*** (0.000)

Note: The table reports summary statistics for the second online pilot experiment regarding perceived characteristics, such as origin, sex, and religious affiliation. The first variables (open-ended questions) are dummy variables that equal 1 if the respondent indicated an origin that matches the respective category. The remaining variables are continuous variables measured on a 7-point Likert scale (ranging from very unlikely to very likely). Column 7 displays a pairwise comparison of differences in means using two-sample t-tests with equal variances. P-values are given in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.17: Second Online Pilot Experiment: Perceived Likelihoods of Origin & Gender by Name

Name	Turkish	German	Saudi Arabian	Moroccan	Italian	Polish	Female	Male
Turkish Female Names								
Aliya Yilmaz	98.25	9.82	34.42	21.90	1.15	1.79	95.03	4.23
Zeynep Yildirim	98.88	10.91	24.00	15.33	0.40	0.74	87.40	15.18
Turkish Male Names								
Muhammed Kaya	93.09	9.38	81.25	68.15	0.90	1.12	0.91	99.38
Ömer Yildiz	99.66	7.36	12.57	9.47	0.75	0.73	2.06	98.41
German Female Names								
Julia Becker	1.89	98.26	1.74	2.94	3.28	10.61	97.94	1.80
Lisa Müller	2.79	97.77	0.00	0.00	1.47	5.84	99.13	0.89
German Male Names								
Maximilian Schäfer	1.76	96.98	0.34	0.35	0.87	3.55	0.87	99.43
Tobias Weber	1.77	97.70	1.01	0.35	3.59	12.72	0.28	99.72

Note: The table reports percentages representing the perceived likelihood of each name being associated with each country and gender measured on a 7-point Likert scale (ranging from very unlikely to very likely). The percentages represent the proportion of participants indicating the respective origin or gender as *very (un)likely* or *quite (un)likely*.

Table A.18: Second Online Pilot Experiment: Treatment Name Manipulation Checks

	(1) Name origin: Turkey	(2) Name origin: Arabic	(3) Name origin: Middle East	(4) Name origin: Turkey
Treatment: Turkish name	0.808***	0.0877***	0.0411***	0.632***
	(0.0121)	(0.00867)	(0.00602)	(0.0593)
Female	0.00129	0.00787	-0.00826	0.00129
	(0.0135)	(0.00957)	(0.00725)	(0.0135)
Age	0.000620	-0.000869*	0.000582	0.000620
	(0.000769)	(0.000488)	(0.000418)	(0.000770)
Student	0.0551**	-0.0148	0.00822	0.0551**
	(0.0233)	(0.0162)	(0.00804)	(0.0234)
Employed	0.0455**	-0.00811	-0.00314	0.0455**
	(0.0220)	(0.0151)	(0.0103)	(0.0221)
Shared apartment	0.00542	-0.00610	0.00437	0.00542
	(0.0129)	(0.00972)	(0.00609)	(0.0129)
Roommates	-0.00267*	0.000726	0.000412	-0.00267*
	(0.00154)	(0.00140)	(0.000623)	(0.00154)
German	0.0692**	-0.0532**	-0.00320	-0.0182
	(0.0282)	(0.0225)	(0.0117)	(0.0130)
Turkish	0.105***	-0.0661***	-0.0172**	-0.00743
	(0.0274)	(0.0142)	(0.00762)	(0.00945)
Migration background	0.0139	-0.0141	-0.00750	-0.00275
	(0.0138)	(0.00974)	(0.00616)	(0.00330)
Treatment × German	-	-	-	0.175***
				(0.0583)
Treatment × Migr. backgr.	-	-	-	0.0332
				(0.0278)
Treatment × Turkish	-	-	-	0.224***
				(0.0535)
Constant	-0.0587	0.0844	-0.0490**	0.0294
	(0.0669)	(0.0553)	(0.0247)	(0.0625)
Additional Survey Controls	Yes	Yes	Yes	Yes
Observations	2,192	2,192	2,192	2,192
R-squared	0.681	0.053	0.026	0.684

Note: The table reports OLS regressions using different dependent variables derived from an open-ended question on the perceived origin of a given name. Responses are coded as dummy variables that equal 1 if the respondent indicated an origin which corresponds to the respective category. Additional survey controls include the time required for completion, a dummy variable indicating whether the survey was fully completed, and an indicator for the first randomly selected treatment condition. Standard errors are clustered at the individual level. Robust standard errors in parenthesis. * p < 0.1, ** p < 0.05, *** p < 0.01.

A.6 Additional Online Survey

To ensure that our social media profiles appear authentic and active – rather than fictitious, inauthentic, or inactive – they must display a network of connections and engagement in the form of followers, subscriptions, comments, and likes. A profile with few followers and limited activity may be perceived as unpopular, disengaged, or suspicious. Consequently, we ask students from our university to follow our fictitious accounts and like a random selection of their posts.

The additional survey was conducted in two waves. In the first wave, which ran between February and March 2020, we recruited $n = 156$ students. Due to the outbreak of Covid-19, the survey was stopped and only resumed in February 2021, continuing until the launch of the actual field experiment in July 2021, at which point an additional $n = 126$ students had participated.

Relying on actual students as friend connections for the social media profiles offers several advantages. First, the demographic characteristics of these followers align closely with the persona depicted, making the profile less likely to be perceived as fictitious. Second, recipients of the profile link – such as the roommate(s) evaluating an applicant for a vacant room – are less likely to doubt the applicant' profile's authenticity if it has a substantial social network and engagement. Finally, this approach reduces the risk that the social networking site will identify the profiles as inauthentic and potentially block them.

Survey Design In a first step, participants are asked to follow a randomly selected subset of the profiles: two out of eight in the case of a public profile and four out of eight in the case of a private profile. Since the list of subscriptions of a public profile is visible to everyone, this approach minimizes the visibility of the experimental profiles to non-participants, thereby reducing the risk of detection. Second, participants are presented with a random selection of images from the respective profiles, where newer images had a higher probability of appearing (see Figure A.9). This simulates the natural growth of a social network over time, as newer posts typically garner more likes. Participants are instructed to remain subscribed to the profiles for the experiment's duration. Additionally, they are asked to evaluate the perceived realism and authenticity of the profiles and to offer any remarks, suggestions, or comments in an open-ended question.

Results A total of $n = 282$ respondents participate in the survey. In terms of perceived realism and authenticity, 91.1% rate the profiles as extremely or very realistic, while 8.2% consider them somewhat realistic and only 0.7% find them only partially realistic. Notably, no respondent perceive them as unrealistic. Therefore, we conclude that our fictitious profiles are sufficiently authentic to support the purposes of our study.

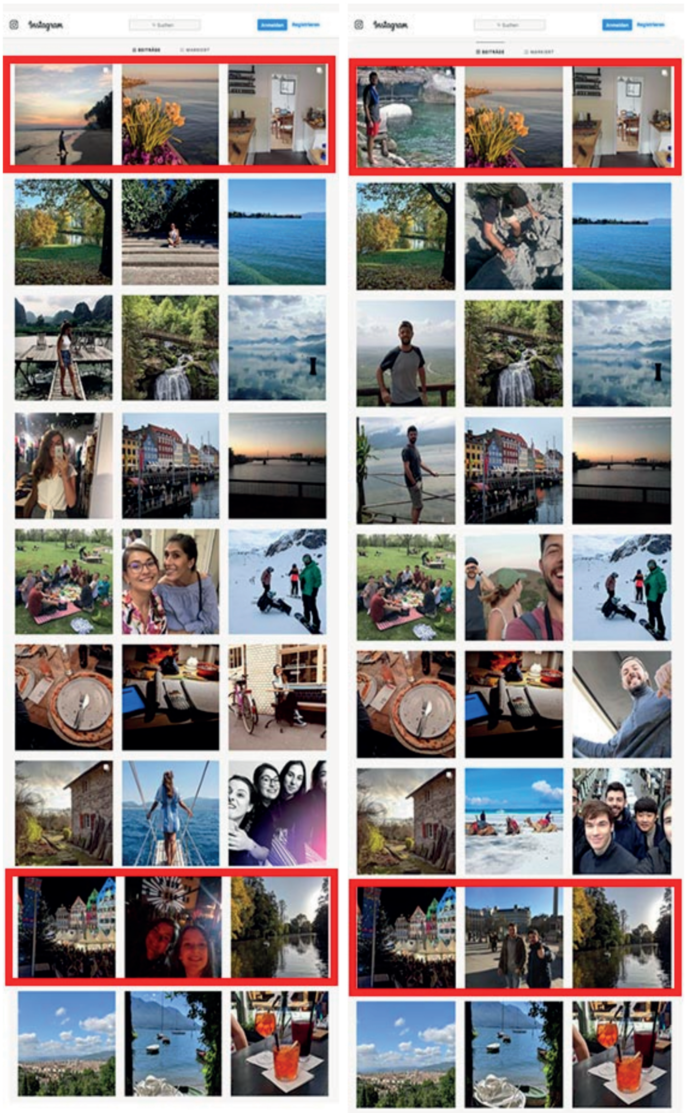
58% ausgefüllt

Ihr Profil: Privat

Bitte **folgen** Sie den vier Profilen. Sie können entweder per Link nacheinander im Browser aufgerufen werden oder auf dem Smartphone in der Instagram-App geöffnet werden. Bitte **liken** Sie anschließend auf jedem der Accounts die Bilder im rot umrandeten Kasten, wie unten dargestellt.

Achtung: Bitte achten Sie darauf, den Fragebogen (dieses Fenster) **nicht** zu schließen!

1. Profil: www.instagram.com/aliya.yilmaz/
2. Profil: www.instagram.com/_muhammed_kaya1/
3. Profil: www.instagram.com/_lisa_mueller1/
4. Profil: www.instagram.com/maximilian_schaefer_/



Zurück

Weiter

Raphael Moritz, Eberhard Karls Universität Tübingen – 2021

Figure A.9: Screenshot of the Follower Task (Additional Online Survey)

Appendix B

Appendix Chapter 3

B.1 Study I: Tables & Figures

Table B.1: Study I: Summary Statistics of Friend Suggestions

		All	**Ethnic Minority**	**Ethnic Majority**	
		Mean/S.D.	Mean/S.D.	Mean/S.D.	Diff.
Private	(1/0)	0.733 (0.442)	0.718 (0.450)	0.748 (0.434)	0.030
Posts	(#)	28.631 (56.938)	34.794 (72.367)	22.593 (34.932)	-12.201***
Followers/subscribers	(#)	577.768 (1004.968)	582.476 (1364.883)	573.157 (419.995)	-9.319
Following/subscriptions	(#)	543.58 (333.959)	567.509 (326.187)	520.142 (340.074)	-47.367*
Mutual connections	(#)	6.362 (5.830)	4.992 (3.817)	7.703 (7.030)	2.711***
New to Instagram	(1/0)	0.028 (0.166)	0.004 (0.062)	0.052 (0.223)	0.048***
Threads account	(1/0)	0.089 (0.284)	0.107 (0.309)	0.071 (0.257)	-0.036*
Story highlights (if public)	(#)	1.106 (2.960)	1.225 (3.099)	0.989 (2.815)	-0.236
Account age	(months)	90.621 (38.044)	94.061 (34.711)	87.252 (40.799)	-6.809**
Former usernames	(#)	1.150 (2.377)	1.109 (2.089)	1.190 (2.629)	0.082
Female	(1/0)	0.708 (0.455)	0.696 (0.460)	0.721 (0.449)	0.025
Business (if public)	(1/0)	0.015 (0.122)	0.013 (0.115)	0.017 (0.129)	0.003
Non-profit (if public)	(1/0)	0.012 (0.110)	0.004 (0.062)	0.021 (0.142)	0.017*
Active story (if public)	(1/0)	0.052 (0.222)	0.069 (0.253)	0.035 (0.185)	-0.033*
Name: Turkish origin	(1/0)	0.009 (0.097)	0.010 (0.097)	0.009 (0.096)	-0.000
Name: German origin	(1/0)	0.696 (0.460)	0.676 (0.468)	0.715 (0.452)	0.038
Bio	(1/0)	0.733 (0.442)	0.710 (0.454)	0.756 (0.430)	0.045
Emojis (in bio)	(#)	1.252 (2.168)	1.404 (2.573)	1.103 (1.667)	-0.301*
Length of bio	(#)	44.422 (42.020)	49.300 (47.196)	39.928 (36.092)	-9.372**
Locations (in bio)	(#)	0.634 (0.885)	0.611 (0.950)	0.657 (0.817)	0.045
Tübingen (in bio)	(1/0)	0.319 (0.466)	0.290 (0.454)	0.347 (0.476)	0.057*
Countries (in bio)	(#)	0.579 (1.073)	0.600 (1.372)	0.558 (0.662)	-0.042
Germany (in bio)	(1/0)	0.421 (0.494)	0.387 (0.487)	0.455 (0.498)	0.069*
Age (in bio)	(years)	22.794 (2.501)	23.033 (2.646)	22.619 (2.386)	-0.414
Student (in bio)	(1/0)	0.180 (0.384)	0.170 (0.376)	0.190 (0.393)	0.021
Motto/saying (in bio)	(1/0)	0.103 (0.304)	0.118 (0.323)	0.090 (0.286)	-0.029
Link (in bio)	(1/0)	0.103 (0.304)	0.086 (0.280)	0.114 (0.318)	0.028
Obs.		1,061	525	536	

Note: The table reports summary statistics for the conditions with ethnic minority names and ethnic majority names. Column 6 displays the results of a two-sample t-test with unequal variances. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.2: Study I: Mean Reaction Rates

	Ethnic Majority	Ethnic Minority	Diff.	Ratio
Panel A: Accept Rates				
All requests	41.04	28.12	-12.92***	1.46
	[402]	[377]	(0.000)	
Without Minority Stereotypes	40.28	30.77	-9.51**	1.31
	[211]	[208]	(0.042)	
With Minority Stereotypes	41.88	24.85	-17.03***	1.69
	[191]	[169]	(0.001)	
Panel B: Re-Follow Rates				
All requests	11.75	11.05	-0.7	1.06
	[536]	[525]	(0.718)	
Without Minority Stereotypes	9.81	14.39	4.58	0.68
	[265]	[264]	(0.106)	
With Minority Stereotypes	13.65	7.66	-5.99**	1.78
	[271]	[261]	(0.035)	
Panel C: Rejection Rates				
All requests	19.40	28.76	9.36***	0.67
	[536]	[525]	(0.000)	
Without Minority Stereotypes	20.75	28.79	8.04**	0.72
	[265]	[264]	(0.033)	
With Minority Stereotypes	18.08	28.74	10.66***	0.63
	[271]	[261]	(0.005)	

Note: The table reports mean reaction rates for different experimental conditions. The numbers in brackets in each cell present the number of requests sent for the given sub-sample and ethnicity. Column 5 shows the p-value for a two-sample Wilcoxon rank-sum (Mann-Whitney) test of proportion testing the null hypothesis that the reaction rates are equal across ethnicities. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.3: Study I: Determinants of Reaction Rates (Interaction Effects)

	(1) Accept	(2) Accept	(3) Re-Follow	(4) Re-Follow
Ethnic Minority	-0.095**	-0.010	0.043	0.048
	(0.046)	(0.048)	(0.027)	(0.042)
Minority Stereotypes	0.013	0.086*	0.037	0.059
	(0.046)	(0.047)	(0.027)	(0.041)
Ethnic Minority × Minority Stereotypes	-0.077	-0.173**	-0.107***	-0.108**
	(0.068)	(0.067)	(0.039)	(0.045)
Posts		0.000		-0.000
		(0.000)		(0.000)
Followers/subscribers		0.000**		0.000
		(0.000)		(0.000)
Following/subscriptions		0.000		0.000
		(0.000)		(0.000)
Mutual connections		-0.003		0.002
		(0.003)		(0.002)
Threads account		-0.001		0.050
		(0.064)		(0.032)
Account age		-0.001**		-0.001**
		(0.001)		(0.000)
Bio		0.073*		0.015
		(0.041)		(0.022)
Female		-0.111***		-0.064***
		(0.038)		(0.020)
Private		0.000		0.072**
		(0.000)		(0.030)
Turkish origin		0.361**		0.269***
		(0.165)		(0.066)
German origin		0.044		-0.013
		(0.039)		(0.021)
Observations	778	774	1,060	1,018
Pseudo R^2	0.016	0.079	0.011	-
Additional Controls	No	Yes	No	Yes

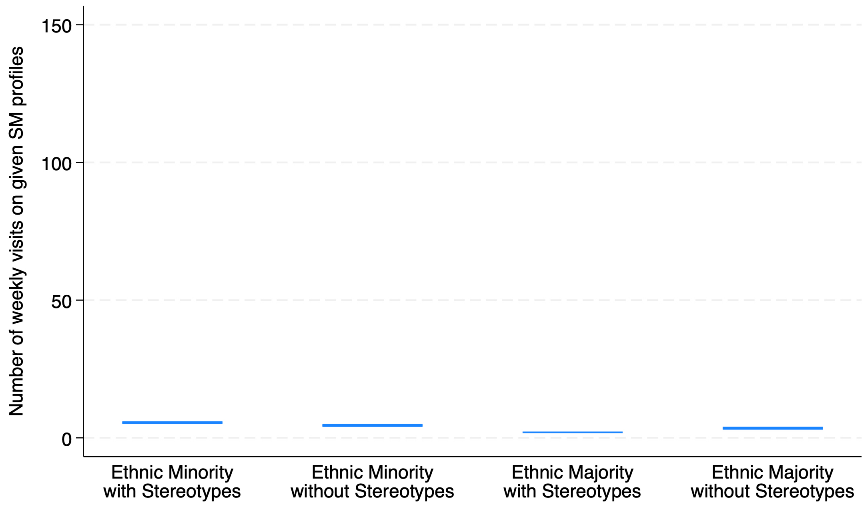
Note: The table reports similar specifications as in Table 3.1, except that it includes the two main effects and their interaction instead of dummies for each treatment condition (see Table 3.1). Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.4: Study I: Weekly Profile Visits per Request (Unique Weekly Visits)

	Mean (S.D.)	Min.	Max.
Panel A: Ethnicity			
Ethnic Minority	1.217 (0.239)	0.893	1.846
Ethnic Majority	1.459 (0.378)	1.104	2.647
Diff.	0.242*** (0.000)		
Panel B: Minority Stereotypes			
With Minority Stereotypes	1.342 (0.343)	1.104	2.647
Without Minority Stereotypes	1.336 (0.336)	0.893	2.500
Diff.	-0.005 (0.788)		
Panel C: With Minority Stereotypes			
Ethnic Minority with Minority Stereotypes	1.268 (0.172)	1.108	1.550
Ethnic Majority with Minority Stereotypes	1.413 (0.439)	1.104	2.647
Diff.	0.145*** (0.000)		
Panel D: Without Minority Stereotypes			
Ethnic Minority without Minority Stereotypes	1.167 (0.283)	0.893	1.846
Ethnic Majority without Minority Stereotypes	1.506 (0.297)	1.245	2.500
Diff.	0.339*** (0.000)		
Total	1.339 (0.339)	0.893	2.647

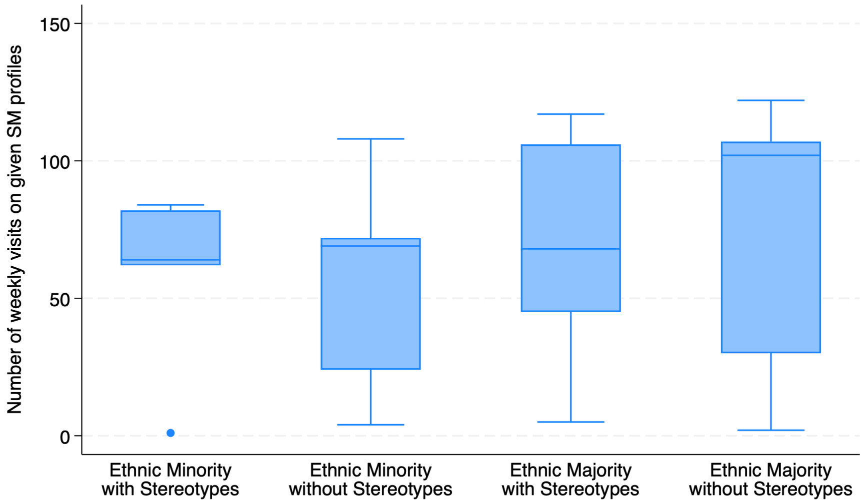
Note: The table reports summary statistics for different experimental conditions on weekly profile visits (unique visits) per request. Differences in means show the results of two-sample t-tests with equal variances. * p < 0.1, ** p < 0.05, *** p < 0.01.

Figure B.1a: Study I: Weekly Profile Visits Before the Experiment



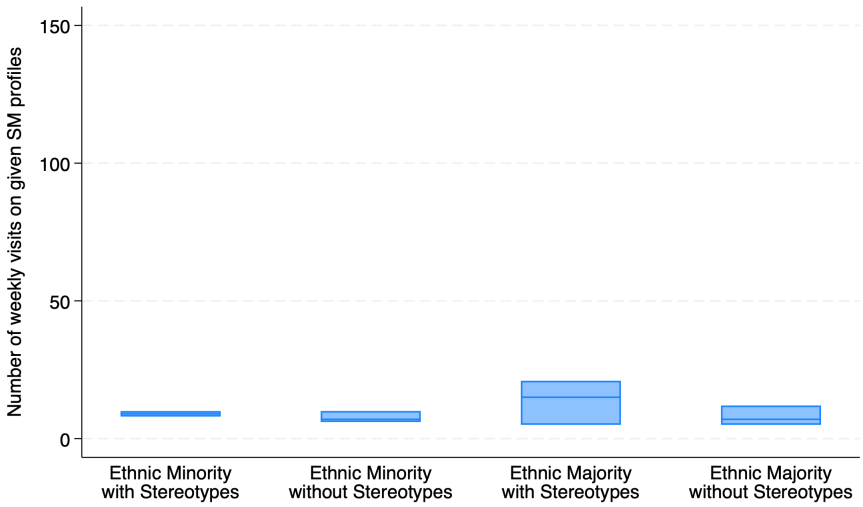
Note: Profile visits are defined as the number of times a profile was accessed (i.e., unique visits) by a user who viewed a post or story (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Figure B.1b: Study I: Weekly Profile Visits During the Experiment



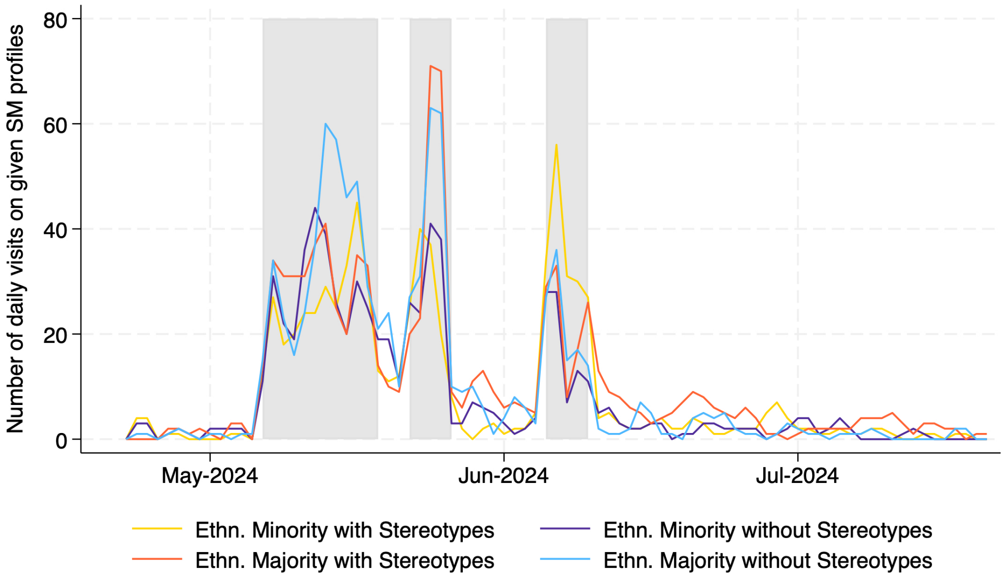
Note: Profile visits are defined as the number of times a profile was accessed (i.e., unique visits) by a user who viewed a post or story (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Figure B.1c: Study I: Weekly Profile Visits After the Experiment



Note: Profile visits are defined as the number of times a profile was accessed (i.e., unique visits) by a user who viewed a post or story (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Figure B.2: Study I: Daily Profile Visits Over the Course of the Experiment



Note: Profile visits are defined as the number of times a profile was accessed (i.e., unique visits) by a user who viewed a post or story (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]). The highlighted sections indicate the days on which the experiment was conducted.

B.1.1 Additional Tables, Figures & Robustness Checks

Table B.5: Study I: Robustness Check: Changes in Mutual Connections

	(1) Accept	(2) Accept	(3) Re-Follow	(4) Re-Follow
Ethnic Majority without Minority Stereotypes (Ref.)	-	-	-	-
Ethnic Majority with Minority Stereotypes	0.120*	0.079	0.076	0.051
	(0.062)	(0.068)	(0.048)	(0.053)
Ethnic Minority without Minority Stereotypes	-0.015	-0.056	0.034	0.034
	(0.060)	(0.066)	(0.049)	(0.054)
Ethnic Minority with Minority Stereotypes	-0.138**	-0.186***	-0.060	-0.042
	(0.062)	(0.066)	(0.055)	(0.060)
Mutual connections	0.002	0.002	0.007**	0.007**
	(0.004)	(0.004)	(0.003)	(0.003)
Changes in mutual connections	-0.008	-0.020***	-0.007	-0.013
	(0.006)	(0.008)	(0.006)	(0.008)
Ethn. Maj. with Min. Stereotypes × Chg. in mutual con.		0.019		0.014
		(0.013)		(0.012)
Ethn. Min. without Min. Stereotypes × Chg. in mutual con.		0.021		0.000
		(0.016)		(0.014)
Ethn. Min. with Min. Stereotypes × Chg. in mutual con.		0.032		-0.037
		(0.020)		(0.046)
Observations	501	501	524	524
Pseudo R^2	0.119	0.124	0.155	0.162
Additional Controls	Yes	Yes	Yes	Yes

Note: The table reports average marginal effects computed from different probit models with accept (columns 1 and 2) and re-follow (columns 3 and 4) as the dependent variable. All specifications include additional control variables (see Table 3.1) and the number of changes in mutual connections (between request and reaction) as an additional robustness check to address a potential source of endogeneity. Columns 2 and 4 include interaction effects of the treatment dummies and the number of changes in mutual connections. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.6: Study I: Neumark Correction for Unobservable Heterogeneity

	(1) Accept	(2) Accept	(3) Re-Follow	(4) Re-Follow
Probit Estimates (Average Marginal Effects)				
Ethnic Majority without Minority Stereotypes (Ref.)	-	-	-	-
Ethnic Majority with Minority Stereotypes	0.013	0.086*	0.037	0.041
	(0.046)	(0.047)	(0.027)	(0.028)
Ethnic Minority without Minority Stereotypes	-0.095**	-0.010	0.043	0.031
	(0.046)	(0.048)	(0.027)	(0.028)
Ethnic Minority with Minority Stereotypes	-0.159***	-0.096*	-0.026	-0.035
	(0.049)	(0.051)	(0.030)	(0.032)
Heteroscedastic Probit Estimates (Average Marginal Effects)				
Ethnic Majority without Minority Stereotypes (Ref.)		-		-
Ethnic Majority with Minority Stereotypes		0.090*		0.059
		(0.053)		(0.041)
Ethnic Minority without Minority Stereotypes		0.030		0.192**
		(0.062)		(0.093)
Ethnic Minority with Minority Stereotypes		-0.051		0.143
		(0.066)		(0.098)
P-value of Wald test for equal variances		0.347		0.025
Observations	778	774	1,060	1,018
Baseline Controls	No	Yes	No	Yes
Additional Controls	No	Yes	No	Yes

Note: The table reports marginal effects computed from different probit and heteroscedastic probit models, with accept as the dependent variable in columns 1 and 2 and re-follow in columns 3 and 4. The baseline controls are presented in Table 3.1. See the notes to Table 3.1 for a list of the additional controls. The p-values in columns 2 and 4 result from a Wald test for equal variances to adjust for potential bias due to unobserved characteristics that may differ between groups. Note that heteroscedastic probit estimates could not be computed for the models in columns 1 and 3. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.7: Study I: Turkish Origin Interactions

	(1) Accept	(2) Accept	(3) Re-Follow	(4) Re-Follow
Ethnic Majority without Minority Stereotypes (Ref.)	-	-	-	-
Ethnic Majority with Minority Stereotypes	0.010	0.083*	0.030	0.035
	(0.046)	(0.047)	(0.027)	(0.027)
Ethnic Minority without Minority Stereotypes	-0.097**	-0.013	0.036	0.023
	(0.046)	(0.049)	(0.027)	(0.028)
Ethnic Minority with Minority Stereotypes	-0.161***	-0.098*	-0.036	-0.042
	(0.049)	(0.051)	(0.030)	(0.032)
Turkish origin	0.088	0.216	-0.632***	-0.645***
	(0.323)	(0.337)	(0.059)	(0.096)
Ethnic Maj. without Min. Stereotypes × Turkish origin	-	-	-	-
Ethnic Maj. with Min. Stereotypes × Turkish origin	0.146	0.171	0.922***	0.816***
	(0.423)	(0.391)	(0.155)	(0.158)
Ethnic Min. without Min. Stereotypes × Turkish origin	-	-	-	-
Ethnic Min. with Min. Stereotypes × Turkish origin	0.161	0.118	0.988***	0.987***
	(0.457)	(0.468)	(0.156)	(0.172)
Observations	777	773	1,058	1,016
Pseudo R^2	0.018	0.077	0.026	0.104
Baseline Controls	No	Yes	No	Yes
Additional Controls	No	Yes	No	Yes

Note: The table reports marginal effects computed from different probit models, with accept as the dependent variable in columns 1 and 2 and re-follow in columns 3 and 4, including interaction terms with the treatment dummies and a dummy that equals one if the treated user's name indicates a Turkish origin. The baseline controls are presented in Table 3.1. See the notes to Table 3.1 for a list of the additional controls. Note that the interaction effect between Ethnic Minority without Minority Stereotypes and Turkish origin is omitted. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.8: Study I: Female Interactions

	(1) Accept	(2) Accept	(3) Re-Follow	(4) Re-Follow
Ethnic Majority without Minority Stereotypes (Ref.)	-	-	-	-
Ethnic Majority with Minority Stereotypes	0.115	0.186**	0.092*	0.121
	(0.099)	(0.094)	(0.050)	(0.078)
Ethnic Minority without Minority Stereotypes	0.019	0.100	0.098**	0.238**
	(0.097)	(0.096)	(0.050)	(0.107)
Ethnic Minority with Minority Stereotypes	-0.135	-0.086	0.006	0.176
	(0.103)	(0.100)	(0.053)	(0.113)
Female	0.011	-0.036	-0.012	-0.030
	(0.082)	(0.080)	(0.046)	(0.072)
Ethnic Maj. without Min. Stereotypes × Female	-	-	-	-
Ethnic Maj. with Min. Stereotypes × Female	-0.131	-0.129	-0.070	-0.087
	(0.112)	(0.107)	(0.060)	(0.091)
Ethnic Min. without Min. Stereotypes × Female	-0.152	-0.141	-0.085	-0.038
	(0.110)	(0.106)	(0.060)	(0.077)
Ethnic Min. with Min. Stereotypes × Female	-0.031	-0.003	-0.058	-0.015
	(0.118)	(0.112)	(0.065)	(0.081)
Observations	775	774	1,021	1,018
Pseudo R^2	0.0230	0.0820	0.0310	-
Baseline Controls	No	Yes	No	Yes
Additional Controls	No	Yes	No	Yes

Note: The table reports marginal effects computed from different probit (columns 1 to 3) and heteroskedastic probit models (column 4), with accept as the dependent variable in columns 1 and 2 and re-follow in columns 3 and 4, including interaction terms with the treatment dummies and a dummy that equals one if the treated user's name indicates a female gender. The baseline controls are presented in Table 3.1. See the notes to Table 3.1 for a list of the additional controls. Robust standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table B.9: Study I: Bio Control Variables

	(1) Accept	(2) Accept	(3) Re-Follow	(4) Re-Follow
Ethnic Majority without Minority Stereotypes (Ref.)	-	-	-	-
Ethnic Majority with Minority Stereotypes	0.086*	0.131**	0.059	0.043
	(0.047)	(0.056)	(0.041)	(0.031)
Ethnic Minority without Minority Stereotypes	-0.010	0.032	0.192**	0.045
	(0.048)	(0.060)	(0.093)	(0.031)
Ethnic Minority with Minority Stereotypes	-0.096*	-0.044	0.143	-0.033
	(0.051)	(0.060)	(0.098)	(0.037)
Length of bio		0.002***		0.001***
		(0.001)		(0.000)
Tübingen in bio		-0.039		0.042
		(0.065)		(0.041)
No. of countries in bio		-0.015		0.016
		(0.020)		(0.012)
Student status in bio		0.002		0.024
		(0.050)		(0.028)
Observations	774	559	1,018	746
Pseudo R^2	0.079	0.079	-	0.143
Baseline Controls	Yes	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes	Yes
Additional Bio Controls	No	Yes	No	Yes

Note: The table reports marginal effects computed from different probit (columns 1, 2, and 4) and heteroskedastic probit models (column 3), with accept as the dependent variable in columns 1 and 2 and re-follow in columns 3 and 4, including control variables on a user's bio. The "bio" section can be found at the top of a user's profile page, where users may add a personal description of up to 150 characters (excluding hyperlinks). The baseline controls are presented in Table 3.1. See the notes to Table 3.1 for a list of the additional controls. Additional (non-tabulated) bio control variables include a dummy variable that equals one if the user indicates residence in Germany, the total number of locations, a dummy variable that equals one if the user adds a motto or saying, and a dummy variable that equals one if the user adds a link to another profile or webpage. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.10: Study I: Determinants of Rejection Rates

	(1) Rejection	(2) Rejection
Ethnic Majority without Minority Stereotypes (Ref.)	-	-
Ethnic Majority with Minority Stereotypes	-0.027	-0.072*
	(0.038)	(0.039)
Ethnic Minority without Minority Stereotypes	0.082**	0.013
	(0.037)	(0.037)
Ethnic Minority with Minority Stereotypes	0.081**	0.070*
	(0.037)	(0.038)
Posts		0.000
		(0.000)
Followers/subscribers		-0.000***
		(0.000)
Following/subscriptions		0.000
		(0.000)
Mutual connections		-0.004
		(0.003)
Threads account		-0.048
		(0.049)
Account age		0.001
		(0.000)
Bio		0.015
		(0.031)
Female		0.056*
		(0.030)
Private		0.285***
		(0.042)
Turkish origin		-0.054
		(0.127)
German origin		0.023
		(0.029)
Observations	1,060	1,018
Pseudo R^2	0.012	0.129
Additional Controls	No	Yes

Note: The table reports the same specifications as in Table 3.1, with rejection as the dependent variable. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.11: Study I: Interactions Proxying a User's Activity (Accept)

	(1) Accept	(2) Accept	(3) Accept	(4) Accept
Ethnic Majority without Minority Stereotypes (Ref.)	-	-	-	-
Ethnic Majority with Minority Stereotypes	0.086	0.051	0.139	0.092*
	(0.056)	(0.082)	(0.096)	(0.048)
Ethnic Minority without Minority Stereotypes	0.022	0.052	-0.029	0.005
	(0.056)	(0.084)	(0.100)	(0.049)
Ethnic Minority with Minority Stereotypes	-0.061	-0.090	-0.054	-0.113**
	(0.057)	(0.090)	(0.109)	(0.053)
Posts	0.001	-0.000	-0.000	0.000
	(0.001)	(0.000)	(0.000)	(0.000)
Ethn. Maj. without Min. Stereotypes × Posts	-			
Ethn. Maj. with Min. Stereotypes × Posts	0.000			
	(0.002)			
Ethn. Min. without Min. Stereotypes × Posts	-0.001			
	(0.001)			
Ethn. Min. with Min. Stereotypes × Posts	-0.002			
	(0.001)			
Length of bio		0.002		
		(0.001)		
Ethn. Maj. without Min. Stereotypes × Length of bio		-		
Ethn. Maj. with Min. Stereotypes × Length of bio		0.003		
		(0.002)		
Ethn. Min. without Min. Stereotypes × Length of bio		-0.000		
		(0.001)		
Ethn. Min. with Min. Stereotypes × Length of bio		0.001		
		(0.002)		
Followers	0.000*	0.000**	0.000**	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Ethn. Maj. without Min. Stereotypes × Followers			-	
Ethn. Maj. with Min. Stereotypes × Followers			-0.000	
			(0.000)	
Ethn. Min. without Min. Stereotypes × Followers			0.000	
			(0.000)	
Ethn. Min. with Min. Stereotypes × Followers			-0.000	
			(0.000)	
Threads account	0.009	0.002	-0.002	0.044
	(0.064)	(0.070)	(0.064)	(0.141)
Ethn. Maj. without Min. Stereotypes × Threads account				-
Ethn. Maj. with Min. Stereotypes × Threads account				-0.139
				(0.200)
Ethn. Min. without Min. Stereotypes × Threads account				-0.201
				(0.181)
Ethn. Min. with Min. Stereotypes × Threads account				0.108
				(0.177)
Observations	774	559	774	774
Pseudo R^2	0.081	0.079	0.080	0.083
Additional Controls	Yes	Yes	Yes	Yes

Note: The table reports similar specifications as in Table 3.1 with accept as the dependent variable, including interactions with treatment conditions and various variables that may serve as proxies for a user's activity. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.12: Study I: Interactions Proxying a User's Activity (Re-Follow)

	(1) Re-Follow	(2) Re-Follow	(3) Re-Follow	(4) Re-Follow
Ethnic Majority without Minority Stereotypes (Ref.)	-	-	-	-
Ethnic Majority with Minority Stereotypes	0.061	0.036	0.120	0.043
	(0.050)	(0.049)	(0.077)	(0.000)
Ethnic Minority without Minority Stereotypes	0.183*	0.006	0.230**	0.121
	(0.098)	(0.050)	(0.110)	(0.000)
Ethnic Minority with Minority Stereotypes	0.140	-0.030	0.164	0.059
	(0.101)	(0.053)	(0.116)	(0.000)
Posts	-0.001	-0.000*	-0.000	-0.000
	(0.001)	(0.000)	(0.000)	(0.000)
Ethn. Maj. without Min. Stereotypes × Posts	-			
Ethn. Maj. with Min. Stereotypes × Posts	-0.000			
	(0.001)			
Ethn. Min. without Min. Stereotypes × Posts	0.001			
	(0.001)			
Ethn. Min. with Min. Stereotypes × Posts	0.000			
	(0.001)			
Length of bio		0.001		
		(0.001)		
Ethn. Maj. without Min. Stereotypes × Length of bio		-		
Ethn. Maj. with Min. Stereotypes × Length of bio		0.000		
		(0.001)		
Ethn. Min. without Min. Stereotypes × Length of bio		0.001		
		(0.001)		
Ethn. Min. with Min. Stereotypes × Length of bio		0.000		
		(0.001)		
Followers	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Ethn. Maj. without Min. Stereotypes × Followers			-	
Ethn. Maj. with Min. Stereotypes × Followers			-0.000	
			(0.000)	
Ethn. Min. without Min. Stereotypes × Followers			-0.000	
			(0.000)	
Ethn. Min. with Min. Stereotypes × Followers			-0.000	
			(0.000)	
Threads account	0.049	0.021	0.051	-1.303
	(0.032)	(0.036)	(0.032)	(0.000)
Ethn. Maj. without Min. Stereotypes × Threads account				-
Ethn. Maj. with Min. Stereotypes × Threads account				1.230
				(0.000)
Ethn. Min. without Min. Stereotypes × Threads account				1.362
				(0.000)
Ethn. Min. with Min. Stereotypes × Threads account				1.413
				(0.000)
Observations	1,018	746	1,018	1,018
Additional Controls	Yes	Yes	Yes	Yes

Note: The table reports the same specifications as in Table B.11 with re-follow as the dependent variable. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.13: Study I: Heterogeneity in Suggestions – Interactions (Accept)

	(1) Accept	(2) Accept	(3) Accept	(4) Accept
Ethnic Majority without Minority Stereotypes (Ref.)	-	-	-	-
Ethnic Majority with Minority Stereotypes	0.093	-0.005	0.090*	0.320**
	(0.102)	(0.072)	(0.047)	(0.131)
Ethnic Minority without Minority Stereotypes	-0.019	-0.063	-0.007	0.115
	(0.106)	(0.077)	(0.048)	(0.135)
Ethnic Minority with Minority Stereotypes	-0.004	-0.149*	-0.094*	0.000
	(0.111)	(0.084)	(0.051)	(0.148)
Following/subscriptions	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Ethn. Maj. without Min. Stereotypes × Following	-			
Ethn. Maj. with Min. Stereotypes × Following	-0.000			
	(0.000)			
Ethn. Min. without Min. Stereotypes × Following	0.000			
	(0.000)			
Ethn. Min. with Min. Stereotypes × Following	-0.000			
	(0.000)			
Mutual connections	-0.003	-0.009*	-0.003	-0.002
	(0.003)	(0.005)	(0.003)	(0.003)
Ethn. Maj. without Min. Stereotypes × Mutual connections		-		
Ethn. Maj. with Min. Stereotypes × Mutual connections		0.011*		
		(0.007)		
Ethn. Min. without Min. Stereotypes × Mutual connections		0.007		
		(0.010)		
Ethn. Min. with Min. Stereotypes × Mutual connections		0.007		
		(0.012)		
New to Instagram	-0.084	-0.070	1.421***	-0.172
	(0.199)	(0.198)	(0.204)	(0.207)
Ethn. Maj. without Min. Stereotypes × New to Instagram			-	
Ethn. Maj. with Min. Stereotypes × New to Instagram			-1.600***	
			(0.273)	
Ethn. Min. without Min. Stereotypes × New to Instagram			0.000	
			(0.000)	
Ethn. Min. with Min. Stereotypes × New to Instagram			0.000	
			(0.000)	
Account age	-0.001**	-0.001**	-0.001**	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Ethn. Maj. without Min. Stereotypes × Account age				-
Ethn. Maj. with Min. Stereotypes × Account age				-0.003*
				(0.001)
Ethn. Min. without Min. Stereotypes × Account age				-0.001
				(0.001)
Ethn. Min. with Min. Stereotypes × Account age				-0.001
				(0.001)
Observations	774	774	774	774
Pseudo R^2	0.0800	0.0810	0.0800	0.0820
Additional Controls	Yes	Yes	Yes	Yes

Note: The table reports similar specifications as in Table 3.1 with accept as the dependent variable, including interactions with treatment conditions to address potential heterogeneity in suggestions. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.14: Study I: Heterogeneity in Suggestions – Interactions (Re-Follow)

	(1) Accept	(2) Accept	(3) Accept	(4) Accept
Ethnic Majority without Minority Stereotypes (Ref.)	-	-	-	-
Ethnic Majority with Minority Stereotypes	0.041	-0.004	0.035	0.160**
	(0.052)	(0.041)	(0.028)	(0.075)
Ethnic Minority without Minority Stereotypes	-0.024	-0.009	0.030	0.118
	(0.053)	(0.040)	(0.028)	(0.078)
Ethnic Minority with Minority Stereotypes	-0.057	-0.057	-0.037	0.049
	(0.058)	(0.052)	(0.032)	(0.086)
Following	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Ethn. Maj. without Min. Stereotypes × Following	-			
Ethn. Maj. with Min. Stereotypes × Following	-0.000			
	(0.000)			
Ethn. Min. without Min. Stereotypes × Following	0.000			
	(0.000)			
Ethn. Min. with Min. Stereotypes × Following	0.000			
	(0.000)			
Mutual connections	0.001	-0.002	0.001	0.002
	(0.002)	(0.003)	(0.002)	(0.002)
Ethn. Maj. without Min. Stereotypes × Mutual connections		-		
Ethn. Maj. with Min. Stereotypes × Mutual connections		0.005		
		(0.004)		
Ethn. Min. without Min. Stereotypes × Mutual connections		0.005		
		(0.005)		
Ethn. Min. with Min. Stereotypes × Mutual connections		0.002		
		(0.008)		
New to Instagram	-0.004	0.018	-0.898***	-0.017
	(0.093)	(0.092)	(0.114)	(0.090)
Ethn. Maj. without Min. Stereotypes × New to Instagram			-	
Ethn. Maj. with Min. Stereotypes × New to Instagram			0.953***	
			(0.144)	
Ethn. Min. without Min. Stereotypes × New to Instagram			0.000	
			(0.000)	
Ethn. Min. with Min. Stereotypes × New to Instagram			0.000	
			(0.000)	
Account age	-0.001**	-0.001**	-0.001**	0.000
	(0.000)	(0.000)	(0.000)	(0.001)
Ethn. Maj. without Min. Stereotypes × Account age				-
Ethn. Maj. with Min. Stereotypes × Account age				-0.001*
				(0.001)
Ethn. Min. without Min. Stereotypes × Account age				-0.001
				(0.001)
Ethn. Min. with Min. Stereotypes × Account age				-0.001
				(0.001)
Observations	1,018	1,018	1,016	1,018
Pseudo R^2	0.106	0.107	0.107	0.107
Additional Controls	Yes	Yes	Yes	Yes

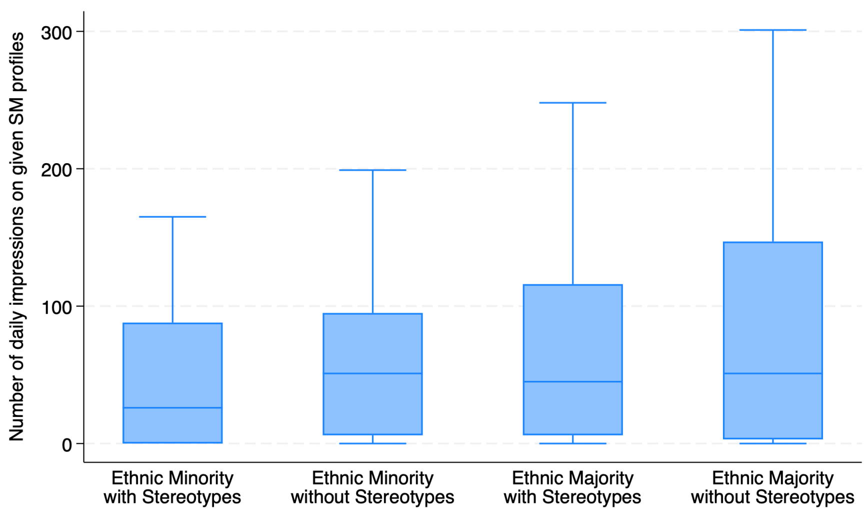
Note: The table reports the same specifications as in Table B.14 with re-follow as the dependent variable. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.15: Study I: Daily Profile Visits (Unique Daily Visits; Approximated)

	Mean (S.D.)	Min.	Max.
Panel A: Ethnicity			
Ethnic Minority	1.646 (0.564)	0.923	4
Ethnic Majority	1.760 (0.728)	0.923	3.929
Diff.	0.115*** (0.004)		
Panel B: Minority Stereotypes			
With Minority Stereotypes	1.711 (0.671)	0.923	3.393
Without Minority Stereotypes	1.697 (0.637)	0.923	4
Diff.	-0.014 (0.729)		
Panel C: With Minority Stereotypes			
Ethnic Minority with Minority Stereotypes	1.824 (0.610)	1.091	4
Ethnic Majority with Minority Stereotypes	1.601 (0.709)	0.923	3.93
Diff.	-0.222*** (0.000)		
Panel D: Without Minority Stereotypes			
Ethnic Minority without Minority Stereotypes	1.469 (0.450)	0.923	3.125
Ethnic Majority without Minority Stereotypes	1.923 (0.712)	1.292	3.393
Diff.	0.453*** (0.000)		
Total	1.704 (0.654)	0.923	4

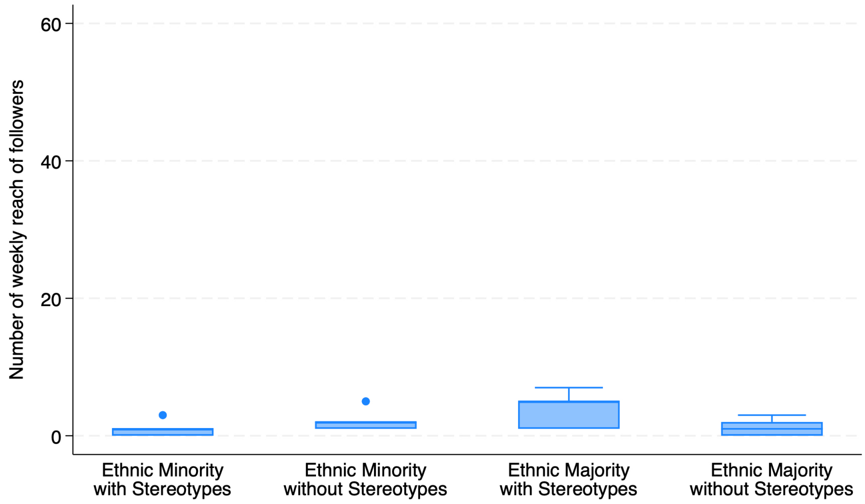
Note: The table reports summary statistics for different experimental conditions on daily profile visits (unique visits) per request. This measure is an approximation, as Instagram provides only limited data on visits (e.g., profile visits from July 3-4 and July 4-5). Therefore, daily profile visits per request is computed as the average of two data points. Differences in means show the results of two-sample t-tests with equal variances. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure B.3: Study I: Daily Impressions



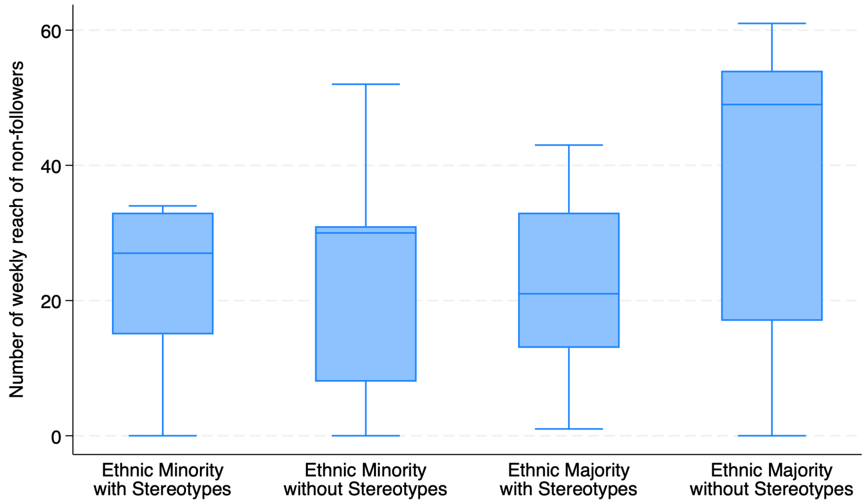
Note: Impressions are defined as the number of times a post or story was on screen (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Figure B.4a: Study I: Weekly Reach of Followers



Note: Reach is defined as the number of unique users that have seen posts or stories at least once. It is different from impressions, which may include multiple views by the same user (see: https://help.instagram.com/788388387972460 and https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Figure B.4b: Study I: Weekly Reach of Non-Followers



Note: Reach is defined as the number of unique users that have seen posts or stories at least once. It is different from impressions, which may include multiple views by the same user (see: https://help.instagram.com/788388387972460 and https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

B.2 Study II: Tables & Figures

Table B.16: Study II: Summary Statistics

		Mean	S.D.	Obs.
Experimental Variables				
Turkish name	(1/0)	0.484	0.500	3,088
Instagram account	(1/0)	0.670	0.470	3,088
Minority Stereotypes	(1/0)	0.333	0.471	3,088
Number of applications per name and week	(#)	171.404	43.903	3,088
Number of SM applications per name and week	(#)	95.993	37.808	3,088
August	(1/0)	0.390	0.488	3,088
September	(1/0)	0.148	0.355	3,088
October	(1/0)	0.462	0.499	3,088
Aachen	(1/0)	0.049	0.216	3,088
Berlin	(1/0)	0.141	0.348	3,088
Bochum	(1/0)	0.023	0.150	3,088
Darmstadt	(1/0)	0.036	0.186	3,088
Dresden	(1/0)	0.043	0.204	3,088
Düsseldorf	(1/0)	0.039	0.193	3,088
Frankfurt am Main	(1/0)	0.057	0.232	3,088
Gießen	(1/0)	0.050	0.218	3,088
Göttingen	(1/0)	0.051	0.220	3,088
Hamburg	(1/0)	0.094	0.291	3,088
Köln	(1/0)	0.072	0.258	3,088
Leipzig	(1/0)	0.058	0.234	3,088
München	(1/0)	0.115	0.319	3,088
Münster	(1/0)	0.038	0.192	3,088
Stuttgart	(1/0)	0.134	0.341	3,088
Response Variables				
Response	(1/0)	0.378	0.485	3,088
Callback	(1/0)	0.259	0.438	3,088
Rejection	(1/0)	0.054	0.226	3,088
Other response	(1/0)	0.064	0.246	3,088
Callback or other response	(1/0)	0.324	0.468	3,088
Number of days between application and response	(days)	2.113	6.340	1,167
Response in same week	(1/0)	0.935	0.247	1,167
Response text includes explicit invitation	(1/0)	0.512	0.500	1,167
Response received outside of platform	(1/0)	0.009	0.097	1,167
Number of characters in response text	(#)	253.440	290.025	1,167
2nd message after initial message	(1/0)	0.023	0.150	1,167
Number of smileys and emojis in response text	(#)	0.248	0.572	1,167
Room & Shared Apartment Characteristics				
Roomsize	(sqm)	17.185	8.157	3,086
Total monthly rent in Euro	(€)	535.364	195.476	3,088
Pictures included	(1/0)	0.975	0.157	3,088
Monthly rent	(€)	457.529	189.153	3,088
Additional monthly costs	(€)	69.139	61.826	3,088
Other monthly costs	(€)	8.697	24.132	3,088
Deposit	(€)	761.155	581.527	3,088
Clearance Payment	(€)	40.545	147.081	3,088
Temporary	(1/0)	0.144	0.351	3,088
Availability (if temporary)	(days)	438.393	393.333	445
Online time	(minutes)	101.252	75.049	3,088

Continued on next page

Table B.16 – continued from previous page

		Mean	S.D.	Obs.
Online viewing possible	(1/0)	0.366	0.482	3,088
Only males wanted	(1/0)	0.079	0.269	3,088
All gender wanted	(1/0)	0.921	0.269	3,088
Age preference for applicants	(1/0)	0.646	0.478	3,088
Number of total roommates	(#)	3.505	1.788	3,088
Apartment size	(sqm)	98.908	56.979	2,336
Apartment size per roommate	(sqm)	30.699	10.325	2,336
Smoking permitted	(1/0)	0.491	0.500	2,301
Number of characters of ad description	(#)	2196.564	1410.815	3,088
Applicant is asked to include social media profile	(1/0)	0.106	0.307	3,088
Ad text states that roommates do not discriminate	(1/0)	0.012	0.109	3,088
Applicant should mention codeword(s)	(1/0)	0.026	0.160	3,088
Number of smileys in ad text	(#)	1.593	2.825	3,088
Roommate Characteristics				
Female roommates	(#)	0.812	0.990	3,088
Male roommates	(#)	1.366	1.421	3,088
Diverse roommates	(#)	0.017	0.167	3,088
Average age	(years)	26.329	5.584	2,321
Roommates speak German	(1/0)	0.854	0.353	3,088
Roommates speak English	(1/0)	0.683	0.465	3,088
Roommates speak Turkish	(1/0)	0.024	0.152	3,088
Roommates speak Arabic	(1/0)	0.018	0.133	3,088
Students	(1/0)	0.659	0.474	3,088
Communally	(1/0)	0.483	0.500	3,088
Non-communally	(1/0)	0.079	0.270	3,088
Males only	(1/0)	0.049	0.216	3,088
Females only	(1/0)	0.048	0.213	3,088
Mixed gender	(1/0)	0.516	0.500	3,088
Young professionals	(1/0)	0.048	0.214	3,088
Employed	(1/0)	0.471	0.499	3,088
Students' hall of residence	(1/0)	0.015	0.122	3,088
Vegetarians/vegans	(1/0)	0.054	0.227	3,088
Cross-generational	(1/0)	0.017	0.131	3,088
Single mother/father	(1/0)	0.006	0.076	3,088
With children	(1/0)	0.009	0.093	3,088
Fraternity	(1/0)	0.077	0.267	3,088
LGBTQIA roommates	(1/0)	0.101	0.301	3,088
Elderly roommates	(1/0)	0.002	0.048	3,088
Disabled roommates	(1/0)	0.021	0.145	3,088
Shared apartment is/will be newly established	(1/0)	0.071	0.257	3,088
Internationals welcome	(1/0)	0.145	0.353	3,088
Advertiser Characteristics				
Account age	(months)	53.369	42.552	3,084
Profile picture	(1/0)	0.479	0.500	3,088
Age	(years)	33.700	13.619	1,155
Female	(1/0)	0.388	0.487	2,861
German origin	(1/0)	0.692	0.462	3,088
Turkish origin	(1/0)	0.011	0.106	3,088
Origin from Arabic country	(1/0)	0.032	0.175	3,088
Origin from a Muslim-majority country	(1/0)	0.069	0.253	3,088
Geographic Variables				
Out of town	(1/0)	0.025	0.157	3,088

Continued on next page

Table B.16 – continued from previous page

		Mean	**S.D.**	**Obs.**
Distance to city center	(kilometers)	4.375	3.419	3,082
Distance to district center	(kilometers)	2.651	2.171	2,691
Distance to main train station	(kilometers)	4.441	3.399	3,085
Bars in 500m radius	(#)	8.624	7.529	3,087
Churches in 1km radius	(#)	11.339	6.421	3,087
Mosques in 1km radius	(#)	1.157	1.998	3,087
Mosques in 3km radius	(#)	5.475	4.633	3,087
University/faculty buildings in 1km radius	(#)	9.611	7.964	3,087
Distance to largest university in city	(kilometers)	5.680	5.241	3,085
Distance to 2nd largest university in city	(kilometers)	5.012	4.017	3,085
Distance to 3rd largest university in city	(kilometers)	6.467	5.653	2,514
Avg. distance to largest two universities in city	(kilometers)	5.346	4.365	3,085

Note: The table presents summary statistics on experimental, response, and geographic variables, as well as room & shared apartment, roommates, and advertiser characteristics.

Table B.17: Study II: Mean Callback Rates

	(1) Ethnic Majority	(2) Ethnic Minority	(3) Diff. if Minority	(4) Ratio Maj/Min
All applications	31.76	19.73	-12.03***	1.61
	[1,593]	[1,495]	(0.000)	
Treatment 1: Without SM	35.01	20.52	-14.49***	1.71
	[517]	[502]	(0.000)	
Treatment 2: SM without Minority Stereotypes	34.69	24.17	-10.52***	1.44
	[562]	[480]	(0.0002)	
Treatment 3: SM with Minority Stereotypes	25.29	14.81	-10.48***	1.71
	[514]	[513]	(0.000)	

Note: The table presents average callback rates for applicants with ethnic majority and ethnic minority names, respectively, depending on the various information treatments. The numbers in brackets indicate the number of applications sent for the given subsample and ethnicity. Column 3 shows the p-values for a two-sample Wilcoxon rank-sum (Mann-Whitney) test of proportion testing the null hypothesis that the callback rates are equal across ethnicities. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.18: Study II: Mean Callback Rates

	(1) Ref. without SM	(2) SM without Min. Stereotypes	(3) SM with Min. Stereotypes	(4) Diff. if (2)	(5) Diff. if (3)	(6) Ratio Ref/(2)	(7) Ratio Ref/(3)
Ethnic Majority	35.01	34.69	25.29	-0.32	-9.72***	1.01	1.38
	[517]	[562]	[514]	(0.914)	(0.001)		
Ethnic Minority	20.52	24.17	14.81	3.65	-5.71**	0.85	1.39
	[502]	[480]	[513]	(0.170)	(0.017)		

Note: The table presents average callback rates for applicants with ethnic majority and ethnic minority names, respectively, depending on the various information treatments. The numbers in brackets in each cell in columns 1-3 indicate the number of applications sent for the given subsample and ethnicity. Columns 4 and 5 show the p-values for a two-sample Wilcoxon rank-sum (Mann-Whitney) test of proportion testing the null hypothesis that the callback rates are equal across ethnicities. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.19: Study II: Probit Models as Split Samples (Average Marginal Effects)

	(1) Ethnic Minority	(2) Ethnic Minority	(3) Ethnic Majority	(4) Ethnic Majority
Treatment 1: Without SM (Ref.)	-	-	-	-
Treatment 2: SM without Minority Stereotypes	0.0353	0.0415	-0.000921	0.00389
	(0.0258)	(0.0258)	(0.0318)	(0.0345)
Treatment 3: SM with Minority Stereotypes	-0.0570***	-0.0265	-0.0995***	-0.0828***
	(0.0187)	(0.0207)	(0.0324)	(0.0308)
Observations	1,495	1,448	1,593	1,544
Month FE	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Room & Shared Apartment Controls	No	Yes	No	Yes
Roommate Controls	No	Yes	No	Yes
Advertiser Controls	No	Yes	No	Yes
Geographic & District Level Demographic Controls	No	Yes	No	Yes

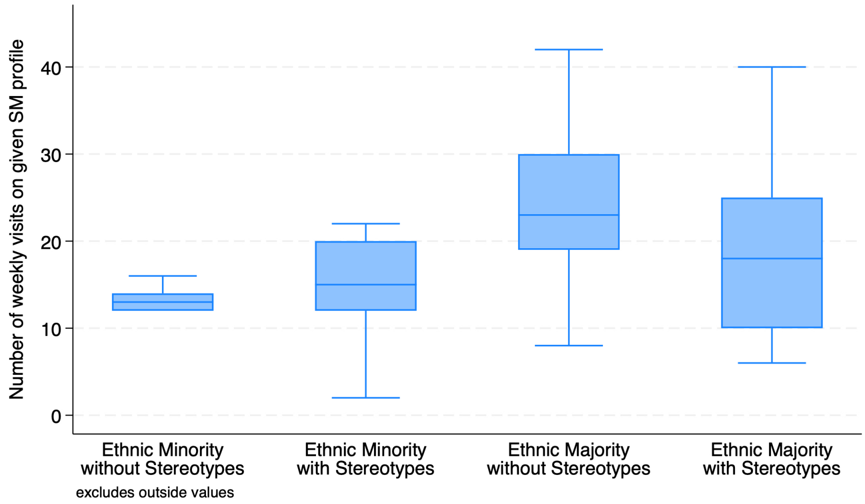
Note: The table reports average marginal effects computed from different probit models with callback (invitation to a viewing) as the dependent variable. Columns 1 and 2 report results for the subsample of ethnic minority applicants, while columns 3 and 4 report results for ethnic majority applicants. Treatment 1 is the reference treatment and is therefore omitted. Columns 2 and 4 report specifications with additional control variables. See the note of Table 3.2 for a description of the control variables. Standard errors (in parentheses) are clustered at the city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.20: Study II: Estimated Average Profile Visits, Impressions, and Reach per Application

	(1) Ethnic Majority	(2) Ethnic Minority	(3) Diff. if Minority	(4) Ratio
Panel A: All Applications				
Profile Visits	0.19	0.13	-0.061***	1.48
	(0.08)	(0.04)	(0.000)	
Impressions	1.47	0.79	-0.675***	1.85
	(0.67)	(0.36)	(0.000)	
Reach of Non-Subscribers	0.10	0.06	-0.037***	1.58
	(0.05)	(0.03)	(0.000)	
Panel B: Without Minority Stereotypes				
Profile Visits	0.20	0.12	-0.083***	1.69
	(0.06)	(0.04)	(0.000)	
Impressions	1.55	0.75	-0.796***	2.06
	(0.60)	(0.41)	(0.000)	
Reach of Non-Subscribers	0.12	0.05	-0.062***	2.14
	(0.04)	(0.03)	(0.000)	
Panel C: With Minority Stereotypes				
Profile Visits	0.17	0.14	-0.039***	1.29
	(0.09)	(0.04)	(0.000)	
Impressions	1.38	0.83	-0.549***	1.66
	(0.73)	(0.30)	(0.000)	
Reach of Non-Subscribers	0.09	0.07	-0.012*	1.16
	(0.05)	(0.02)	(0.082)	

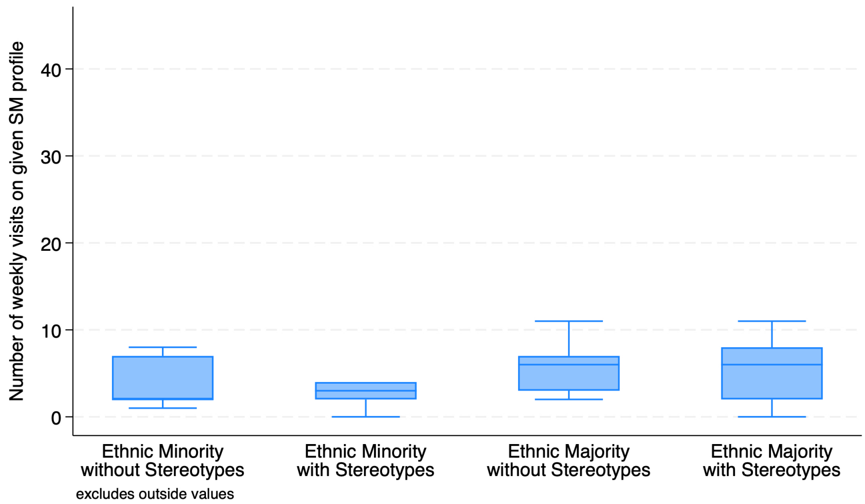
Note: The table reports means and standard deviations (in parentheses) of calculating the estimated average profile visits, impressions, and reach of non-subscribers per application by computing the number of visits, impressions, and reach of a given social media (SM) profile in a given week divided by the number of applications sent per name in the same week in the SM conditions (Treatments including a SM profile) in columns 1 and 2. Column 3 shows the p-values for a two-sample Wilcoxon rank-sum (Mann-Whitney) test of proportion testing the null hypothesis that the visits, impressions, and reach are equal across ethnicities. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure B.5a: Study II: Weekly Profile Visits Over the Course of the Experiment



Note: Profile visits are defined as the number of times a profile was accessed (i.e., unique visits) by a user who viewed a post or story (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Figure B.5b: Study II: Weekly Profile Visits Before and After the Experiment



Note: Profile visits are defined as the number of times a profile was accessed (i.e., unique visits) by a user who viewed a post or story (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Table B.21: Study II: OLS Models on Social Media Variables

	(1) Weekly Profile Visits	(2) Weekly Reach of Non-Subscribers	(3) Weekly Impressions
Ethnic Minority	-56.34***	-43.77***	-498.1***
	(3.152)	(2.603)	(34.98)
Treatment 2: SM without Minority Stereotypes (Ref.)	-	-	-
Treatment 3: SM with Minority Stereotypes	-27.66***	-13.16***	-163.7***
	(2.985)	(1.381)	(23.23)
Ethnic Minority × Treatment 2 (Ref.)	-	-	-
Ethnic Minority × Treatment 3	78.78***	69.07***	481.5***
	(4.397)	(4.536)	(58.50)
Number of roommates	0.0397	0.0361	-0.377
	(0.0609)	(0.0563)	(0.790)
Monthly rent	-0.000629	-0.000427	-0.00182
	(0.000730)	(0.000562)	(0.0107)
Temporary	-1.133***	-0.749**	-12.26***
	(0.306)	(0.269)	(3.179)
Female roommates	-0.156	-0.123	-0.756
	(0.0996)	(0.0912)	(1.451)
Students	0.287	0.0638	3.451
	(0.335)	(0.250)	(2.969)
Appl. should include SM profile	0.675**	0.610***	10.93**
	(0.309)	(0.202)	(3.748)
Likes on all posts	-0.417***	-0.495***	-1.862**
	(0.0389)	(0.0437)	(0.677)
Followers/subscribers	-0.0530	0.393***	3.112**
	(0.119)	(0.0833)	(1.165)
Following/subscriptions	2.796***	2.568***	19.38***
	(0.143)	(0.171)	(2.186)
Constant	-204.6***	-172.9***	-3,317***
	(51.18)	(42.56)	(564.1)
Observations	2,069	2,069	2,069
R-squared	0.428	0.430	0.380
Month FE	Yes	Yes	Yes
City FE	Yes	Yes	Yes

Note: The table reports different OLS models with weekly profile visits, weekly reach of non-subscribers, and weekly impressions as dependent variables and various social media and treatment variables as independent variables. Standard errors (in parentheses) are clustered at the city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.2.1 Additional Tables, Figures & Robustness Checks

Table B.22: Study II: Summary Statistics on District Population Variables

	Mean	**S.D.**	**Obs.**
Population 2017	125,887	115,050	3,008
Population 2018	126,916	116,099	3,008
Population 2019	127,641	116,565	3,008
Population 2020	127,537	116,560	3,008
Population 2021	127,353	116,602	3,008
Population change 2017-2021 (annual growth rate)	0.003	0.008	3,008
Share of female population 2017	0.501	0.019	3,008
Share of female population 2018	0.501	0.018	3,008
Share of female population 2019	0.496	0.041	3,008
Share of female population 2020	0.503	0.016	3,008
Share of female population 2021	0.501	0.018	3,008
Foreign population 2017	25,793	26,589	3,008
Foreign population 2018	26,545	27,602	3,008
Foreign population 2019	27,184	28,104	3,008
Foreign population 2020	27,304	28,070	3,008
Foreign population 2021	27,865	28,510	3,008
Foreign population change 2017-2021 (annual growth rate)	0.016	0.032	3,008
Share of foreign population 2017	0.214	0.085	3,008
Share of foreign population 2018	0.213	0.084	3,008
Share of foreign population 2019	0.216	0.083	3,008
Share of foreign population 2020	0.216	0.082	3,008
Share of foreign population 2021	0.222	0.082	3,008
Turkish population 2017	4,565	7,883	3,008
Turkish population 2018	4,553	7,961	3,008
Turkish population 2019	4,546	7,909	3,008
Turkish population 2020	4,502	7,809	3,008
Turkish population 2021	4,484	7,745	3,008
Turkish population change 2017-2021 (annual growth rate)	0.002	0.052	3,007
Share of Turkish population 2017	0.029	0.024	3,008
Share of Turkish population 2018	0.027	0.024	3,008
Share of Turkish population 2019	0.027	0.023	3,008
Share of Turkish population 2020	0.027	0.023	3,008
Share of Turkish population 2021	0.027	0.023	3,008
Population density	5,829	5,875	2,855
Households	74,428	64,544	2,855
Apartments	61,507	51,636	2,593
Residential buildings	13,068	13,872	2,165
Rental price (€ per sqm)	12.185	3.963	2,199
Unemployment rate	0.043	0.032	1,311
Average age	40.919	2.347	2,900
German population age 18-30	17,644	13,742	2,344
Foreign population age 18-30	6,513	6,570	2,746
Population age 18-30	23,729	20,069	2,894
Share of migrants	0.260	0.114	2,746
Students in city	87,679	49,559	3,008
Rent control ("Mietpreisbremse") in city	0.833	0.373	3,008

Note: The table shows summary statistics on district-level population and socio-demographic variables.

Table B.23: Study II: Summary Statistics on Election Results

	Mean	**S.D.**
Municipal Elections		
Eligible voters	894,952	843,812
Voters	485,116	499,601
Voter turnout	52.120	6.671
Share of votes for SPD	29.095	10.651
Change (pp.) in the share of votes for SPD	-7.132	8.802
Share of votes for CDU	29.210	8.266
Change (pp.) in the share of votes for CDU	-4.114	10.523
Share of votes for Gruene	24.859	9.311
Change (pp.) in the share of votes for Gruene	6.937	10.289
Share of votes for Linke	8.594	5.564
Change (pp.) in the share of votes for Linke	-1.283	1.758
Share of votes for AFD	6.032	3.384
Change (pp.) in the share of votes for AFD	-0.444	2.872
Share of votes for FDP	3.890	2.748
Change (pp.) in the share of votes for FDP	0.943	3.902
Share of votes for NPD	0.222	0.514
Change (pp.) in the share of votes for NPD	-0.282	0.252
State Elections		
Eligible voters	794,906	752,643
Voters	505,416	474,517
Voter turnout	63.461	4.043
Share of (primary) votes for SPD	19.739	8.201
Change (pp.) in the share of (primary) votes for SPD	-5.384	3.455
Share of (primary) votes for CDU	25.902	5.016
Change (pp.) in the share of (primary) votes for CDU	-0.040	5.904
Share of (primary) votes for Gruene	27.240	7.113
Change (pp.) in the share of (primary) votes for Gruene	8.117	8.327
Share of (primary) votes for Linke	7.533	4.943
Change (pp.) in the share of (primary) votes for Linke	-2.217	2.079
Share of (primary) votes for AFD	8.471	4.963
Change (pp.) in the share of (primary) votes for AFD	0.130	5.655
Share of (primary) votes for FDP	5.848	2.190
Change (pp.) in the share of (primary) votes for FDP	-1.032	1.719
Share of (primary) votes for NPD	0.000	0.000
Change (pp.) in the share of (primary) votes for NPD	-0.300	0.000
Share of (secondary) votes for SPD	20.020	8.763
Change (pp.) in the share of (secondary) votes for SPD	-4.344	2.152
Share of (secondary) votes for CDU	25.357	5.739
Change (pp.) in the share of (secondary) votes for CDU	1.556	5.653
Share of (secondary) votes for Gruene	24.666	5.355
Change (pp.) in the share of (secondary) votes for Gruene	7.381	8.074
Share of (secondary) votes for Linke	7.074	4.329
Change (pp.) in the share of (secondary) votes for Linke	-3.716	1.930
Share of (secondary) votes for AFD	7.757	5.366
Change (pp.) in the share of (secondary) votes for AFD	-0.024	4.789
Share of (secondary) votes for FDP	5.491	1.051
Change (pp.) in the share of (secondary) votes for FDP	-3.154	3.014
Share of (secondary) votes for NPD	0.184	0.099
Change (pp.) in the share of (secondary) votes for NPD	-1.551	1.237
Bundestag (Federal) Elections 2021		
Eligible voters	799,482	761,111

Continued on next page

Table B.23 – continued from previous page

	Mean	**S.D.**
Voters	594,882	530,211
Voter turnout	76.595	3.770
Share of (primary) votes for SPD	23.534	5.635
Change (pp.) in the share of (primary) votes for SPD	-0.978	3.370
Share of (primary) votes for CDU	22.430	3.353
Change (pp.) in the share of (primary) votes for CDU	-7.541	2.076
Share of (primary) votes for Gruene	25.140	6.085
Change (pp.) in the share of (primary) votes for Gruene	11.369	4.307
Share of (primary) votes for Linke	7.815	4.761
Change (pp.) in the share of (primary) votes for Linke	-3.713	1.374
Share of (primary) votes for AFD	6.536	3.750
Change (pp.) in the share of (primary) votes for AFD	-2.617	1.548
Share of (primary) votes for FDP	8.735	1.783
Change (pp.) in the share of (primary) votes for FDP	1.252	1.076
Share of (primary) votes for NPD	0.100	0.000
Change (pp.) in the share of (primary) votes for NPD	-0.100	0.000
Share of (secondary) votes for SPD	23.008	3.502
Change (pp.) in the share of (secondary) votes for SPD	3.654	2.443
Share of (secondary) votes for CDU	18.747	3.267
Change (pp.) in the share of (secondary) votes for CDU	-7.776	1.921
Share of (secondary) votes for Gruene	24.811	3.713
Change (pp.) in the share of (secondary) votes for Gruene	10.717	2.615
Share of (secondary) votes for Linke	7.374	2.901
Change (pp.) in the share of (secondary) votes for Linke	-5.340	1.146
Share of (secondary) votes for AFD	6.662	3.559
Change (pp.) in the share of (secondary) votes for AFD	-3.154	0.796
Share of (secondary) votes for FDP	11.699	2.573
Change (pp.) in the share of (secondary) votes for FDP	-0.434	1.558
Share of (secondary) votes for NPD	0.066	0.064
Change (pp.) in the share of (secondary) votes for NPD	-0.121	0.091
European Elections 2024		
Eligible voters	740,031	785,537
Voters	478,636	486,549
Voter turnout	66.826	3.669
Share of votes for SPD	13.837	3.111
Change (pp.) in the share of votes for SPD	0.847	5.165
Share of votes for CDU	23.065	6.342
Change (pp.) in the share of votes for CDU	2.350	3.891
Share of votes for Gruene	19.977	4.055
Change (pp.) in the share of votes for Gruene	-6.596	7.712
Share of votes for Linke	4.653	2.285
Change (pp.) in the share of votes for Linke	-2.099	2.419
Share of votes for AFD	9.917	4.004
Change (pp.) in the share of votes for AFD	1.927	1.233
Share of votes for FDP	6.269	1.927
Change (pp.) in the share of votes for FDP	0.583	1.448
Share of votes for NPD/Heimat	0.036	0.057
Change (pp.) in the share of votes for NPD/Heimat	-0.066	0.078

Note: The table presents summary statistics on election results for municipal, state, federal, and European elections at the city level. Note that vote shares are given as percentages, not decimal shares.

Table B.24: Study II: Summary Statistics on Social Media Variables

	Mean	S.D.	Mean	S.D.
	Ethnic Minority with Minority Stereotypes		Ethnic Majority with Minority Stereotypes	
Posts	32	0	32	0
Stories	15	0	15	0
Followers/subscribers	226	3	211	3
Following/subscriptions	264	0	262	1
Likes on all posts	1225	6	1145	7
Reach	8	2	12	6
Reach of subscribers	0	0	1	1
Reach of non-subscribers	8	2	10	6
Impressions	89	33	158	96
Profile visits	15	5	19	10
Reach of first post	268	12	210	19
Reach of subscribers of first post	159	1	141	0
Reach of non-subscribers of first post	110	11	69	19
Impressions of first post	328	14	244	23
Reel playbacks	137	8	139	12
Obs.	513		514	
	Ethnic Minority without Minority Stereotypes		Ethnic Majority without Minority Stereotypes	
Posts	27	0	27	0
Stories	15	0	15	0
Followers/subscribers	233	3	229	2
Following/subscriptions	274	0	249	1
Likes on all posts	1172	8	1109	2
Reach	7	3	15	6
Reach of subscribers	1	1	0	1
Reach of non-subscribers	6	3	15	6
Impressions	81	40	192	88
Profile visits	12	3	25	10
Obs.	480		562	

Note: The table shows summary statistics on the social media variables on a weekly basis.

Table B.25: Study II: Number of Applications, Responses and Type of Response

	Applications	Responses	Callbacks	Others	Rejections
Ethnic Majority					
Treatment 1: Without SM	517	245	181	40	23
Treatment 2: SM without Minority Stereotypes	562	265	195	36	34
Treatment 3: SM with Minority Stereotypes	514	186	130	37	19
Total	1,593	696	506	113	76
Ethnic Minority					
Treatment 1: Without SM	502	155	103	23	29
Treatment 2: SM without Minority Stereotypes	480	173	116	28	29
Treatment 3: SM with Minority Stereotypes	513	143	76	35	32
Total	1,495	471	295	86	90
Total	3,088	1,167	801	199	166

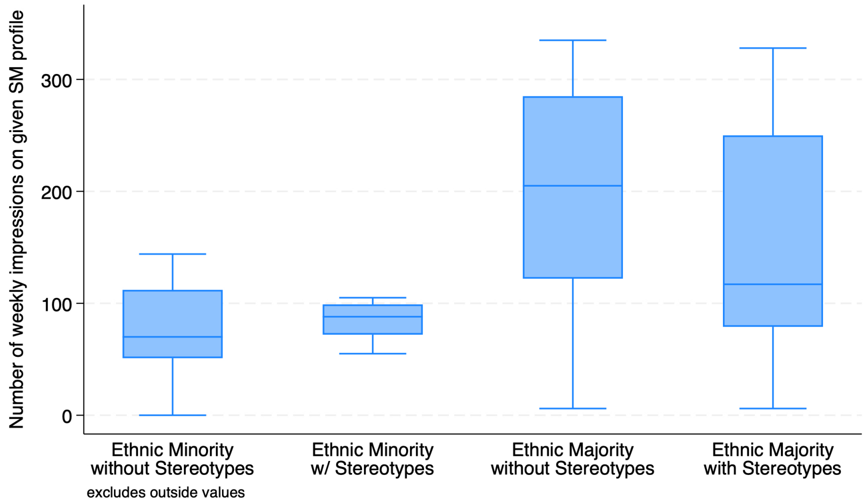
Note: The table reports the absolute numbers of applications, responses, callbacks, other responses, and rejections by ethnicity and information treatment.

Table B.26: Study II: Probit Models with Interaction Effects (Average Marginal Effects) – Rejections

Rejection	(1)	(2)	(3)	(4)
Ethnic Minority	0.0119	0.0115	0.0125	0.0126
	(0.00862)	(0.0127)	(0.00830)	(0.0131)
Treatment 1: Without SM (Ref.)	-	-	-	-
Treatment 2: SM without Minority Stereotypes	0.00993	0.0149	0.00626	0.0111
	(0.0112)	(0.0113)	(0.0117)	(0.0101)
Treatment 3: SM with Minority Stereotypes	0.00177	-0.00580	-0.00390	-0.0108
	(0.0138)	(0.0163)	(0.0129)	(0.0183)
Ethnic Minority × Treatment 1 (Ref.)		-		-
Ethnic Minority × Treatment 2		-0.0104		-0.0120
		(0.0223)		(0.0234)
Ethnic Minority × Treatment 3		0.0134		0.0140
		(0.0263)		(0.0274)
Observations	3,088	3,088	2,998	2,998
Month FE	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Room & Shared Apartment Controls	No	No	Yes	Yes
Roommate Controls	No	No	Yes	Yes
Advertiser Controls	No	No	Yes	Yes
Geographic & District Level Demographic Controls	No	No	Yes	Yes

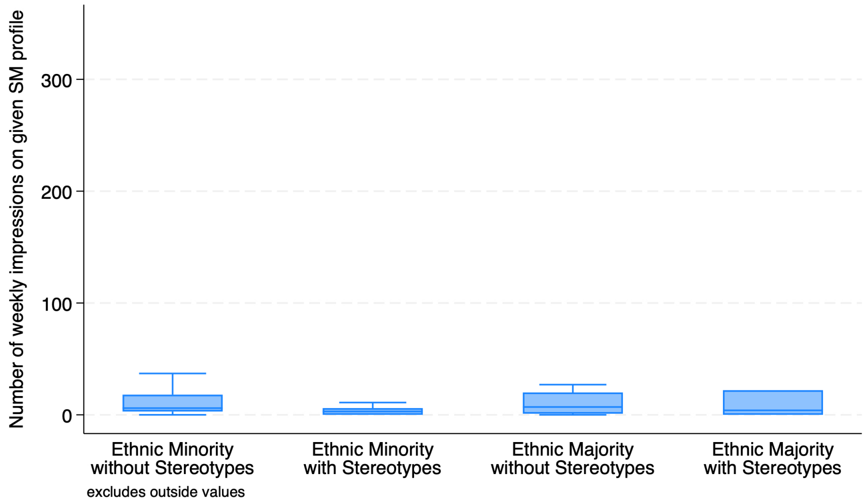
Note: The table reports average marginal effects computed from different probit models with rejection as the dependent variable. Columns 1 and 2 report results from regular probit models, columns 3 and 4 from heteroscedastic probit models. Columns 2 and 4 report additional interaction effects of ethnic minority and treatment variables. Treatment 1 is the reference treatment and is therefore omitted. Columns 3 and 4 report specifications with additional control variables. See the note of Table 3.2 for a description of the control variables. Standard errors (in parentheses) are clustered at the city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure B.6a: Study II: Weekly Impressions Over the Course of the Experiment



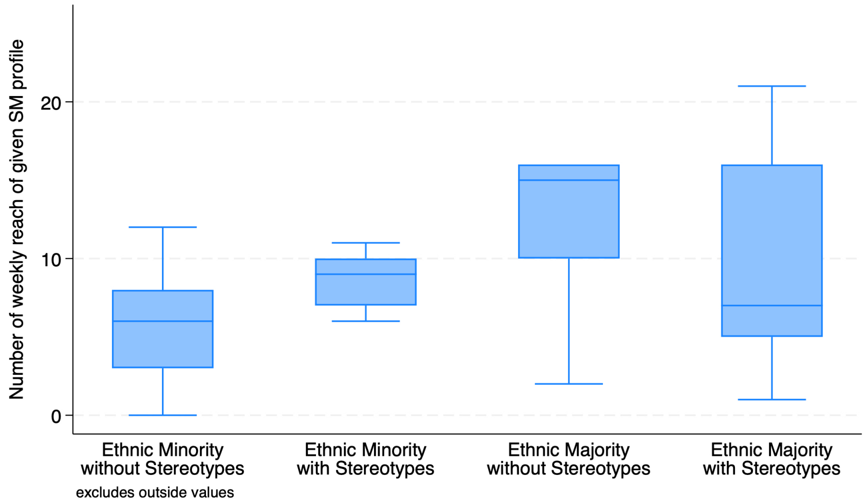
Note: Impressions are defined as the number of times a post or story was on screen (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Figure B.6b: Study II: Weekly Impressions Before and After the Experiment



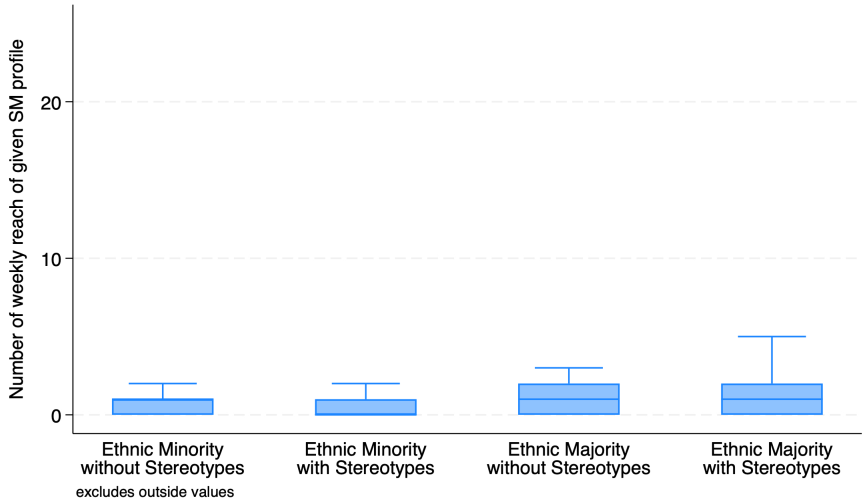
Note: Impressions are defined as the number of times a post or story was on screen (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Figure B.7a: Study II: Weekly Reach Over the Course of the Experiment



Note: Reach is defined as the number of unique users that have seen posts or stories at least once. It is different from impressions, which may include multiple views by the same user (see: https://help.instagram.com/788388387972460 and https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Figure B.7b: Study II: Weekly Reach Before and After the Experiment



Note: Reach is defined as the number of unique users that have seen posts or stories at least once. It is different from impressions, which may include multiple views by the same user (see: https://help.instagram.com/788388387972460 and https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Table B.27: Study II: Neumark Correction for Unobservable Heterogeneity

Callback	(1)	(2)	(3)
Probit Estimates (Average Marginal Effects)			
Ethnic Minority	-0.140***	-0.141***	-0.145***
	(0.0235)	(0.0204)	(0.0166)
Treatment 1: Without SM (Ref.)	-	-	-
Treatment 2: SM without Minority Stereotypes	-0.00436	-0.00817	-0.000242
	(0.0288)	(0.0267)	(0.0290)
Treatment 3: SM with Minority Stereotypes	-0.0909**	-0.0867**	-0.0782**
	(0.0287)	(0.0277)	(0.0255)
Ethnic Minority × Treatment 1: Without SM (Ref.)	-	-	-
Ethnic Minority × Treatment 2: SM without Minority Stereotypes	0.0449	0.0501	0.0479
	(0.0336)	(0.0304)	(0.0282)
Ethnic Minority × Treatment 3: SM with Minority Stereotypes	0.0253	0.0308	0.0380
	(0.0240)	(0.0284)	(0.0217)
Heteroscedastic Probit Estimates (Average Marginal Effects)			
Ethnic Minority	-0.140***	-0.140***	-0.144***
	(0.0236)	(0.0208)	(0.0179)
Treatment 1: Without SM (Ref.)	-	-	-
Treatment 2: SM without Minority Stereotypes	-0.00432	-0.00849	-0.00130
	(0.0285)	(0.0275)	(0.0313)
Treatment 3: SM with Minority Stereotypes	-0.0905***	-0.0897**	-0.0862**
	(0.0270)	(0.0277)	(0.0271)
Ethnic Minority × Treatment 1: Without SM (Ref.)	-	-	-
Ethnic Minority × Treatment 2: SM without Minority Stereotypes	0.0451	0.0486	0.0447
	(0.0332)	(0.0306)	(0.0292)
Ethnic Minority × Treatment 3: SM with Minority Stereotypes	0.0245	0.0370	0.0555*
	(0.0223)	(0.0278)	(0.0220)
P-value of Wald test for equal variances	0.9443	0.5608	0.0159
Observations	3,088	2,998	2,992
Pseudo R-squared	0.0749	0.1202	0.1872
Month FE	Yes	Yes	Yes
City FE	Yes	Yes	Yes
Roommate controls	No	Yes	Yes
Geographic & district-level demographic controls	No	Yes	Yes
Room & shared apartment controls	No	No	Yes
Advertiser controls	No	No	Yes

Note: The table reports marginal effects computed from different probit and heteroscedastic probit models, with callback (invitation to a viewing) as the dependent variable. Treatment 1 is the reference treatment and is therefore omitted. The models in column 1 include only baseline controls, while in column 2 they include additional controls for roommate characteristics and geographic & district-level demographic variables in column 2. In column 3, they also include room & shared apartment and advertiser controls. See the note of Table 3.2 for a description of the control variables. The p-values in columns 1 to 3 result from a Wald test for equal variances to adjust for potential bias due to unobserved characteristics that may differ between groups. The model in column 3 provides significant evidence that the variance of unobservables differs between ethnic minority and majority applicants when room & shared apartment and advertiser controls are included. Standard errors (in parentheses) are clustered at the city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.28: Study II: Election Results: Probit Models with Interaction Effects (Average Marginal Effects) – Callback

Callback	(1) Municipal AFD	(2) Municipal NPD	(3) State AFD	(4) State NPD	(5) Federal AFD	(6) Federal NPD	(7) European AFD	(8) European NPD
Ethnic Minority	-0.0694***	-0.140***	-0.0905***	-0.105***	-0.0951***	-0.0803***	-0.0845***	-0.102***
	(0.0230)	(0.0340)	(0.0208)	(0.0103)	(0.0280)	(0.0120)	(0.0312)	(0.0199)
Treatment 1 (Ref.)	-	-	-	-	-	-	-	-
Treatment 2	0.0414	0.0735***	0.0216	0.0296	0.0218	0.0106	0.0225	0.0222
	(0.0272)	(0.0220)	(0.0281)	(0.0382)	(0.0246)	(0.0278)	(0.0249)	(0.0249)
Treatment 3	-0.0367	-0.0302**	-0.0622***	-0.0191	-0.0622***	-0.0662**	-0.0625***	-0.0626***
	(0.0244)	(0.0141)	(0.0241)	(0.0234)	(0.0223)	(0.0257)	(0.0225)	(0.0225)
Voter Turnout	-0.00340	-0.313***	-0.0114**	0.0176***	-0.0120*	-0.0262***	-0.0146**	-0.0145**
	(0.00454)	(0.0129)	(0.00555)	(0.00357)	(0.00657)	(0.00998)	(0.00681)	(0.00696)
Municipal Elections								
AFD Share	0.00137							
	(0.00393)							
Ethn. Min. × AFD	-0.00563							
	(0.00379)							
NPD Share		-1.428***						
		(0.0332)						
Ethn. Min. × NPD		-0.00906						
		(0.0371)						
State Elections								
AFD Share			0.0111***					
			(0.00287)					
Ethn. Min. × AFD			-0.00326*					
			(0.00193)					
NPD Share				2.715***				
				(0.392)				
Ethn. Min. × NPD				0.0628				
				(0.0418)				
Federal (Bundestag) Elections 2021								
AFD Share					-0.00245			
					(0.00422)			
Ethn. Min. × AFD					-0.00256			
					(0.00217)			
NPD Share						-0.295		
						(0.353)		
Ethn. Min. × NPD						-0.548***		
						(0.209)		
European Elections 2024								
AFD Share							0.00237	
							(0.00327)	
Ethn. Min. × AFD							-0.00268	
							(0.00201)	
NPD Share								0.153
								(0.289)
Ethn. Min. × NPD								-0.255
								(0.172)
Observations	1,990	774	2,590	734	2,992	2,567	2,992	2,992
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
All controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: The table reports average marginal effects computed from different probit models with callback (invitation to a viewing) as the dependent variable. Columns 1, 3, and 5 compute shares of votes for the right-wing party "AFD" ("Alternative für Deutschland"), while columns 2, 4, and 6 compute shares of votes for the right-wing extremist party "NPD" ("Nationaldemokratische Partei Deutschlands", since 2023: "Die Heimat"), including interactions with an ethnic minority dummy. Columns 1 and 2 estimate the effects for municipal elections (2019-2023), columns 3 and 4 for state elections, columns 5 and 6 the federal (Bundestag) elections in 2021, and columns 7 and 8 for the European elections in 2024. The shares of votes for state and federal elections consist of secondary votes only, as the share of secondary votes determines the respective party's representation in parliament (see: https://www.bmi.bund.de/EN/topics/constitution/electoral-law/bundestag-elections/bundestag-elections-node.html [Retrieved: March 3, 2025]). Treatment 1 is the reference treatment and is therefore omitted. All controls refers to room & shared apartment, roommate, advertiser, and geographic & district level demographic controls (see the note of Table 3.2 for a description of the control variables). Standard errors (in parentheses) are clustered at the city level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.3 Study II: Response Classification

The main outcome of the experiment is callbacks to measure the advertiser's level of interest in the applicant. All responses from advertisers and roommates are classified into one of three main categories: callbacks, rejections, or 'other' responses. The majority of 'other' responses pertain to additional queries or requests for additional information about the applicant. Callbacks include an invitation to a viewing, meeting, or call. The following provides a list of examples of advertiser responses.

Callback: All callbacks that include a clear interest in offering the opportunity to view the room, and/or to meet the potential roommates, or to have a telephone interview are classified as callbacks. This may be indicated by the mention of a specific date (e.g., "Friday, 4 October at 7pm"), by asking for the applicant's availability for a meeting or call, or by expressing a willingness to meet the applicant (e.g., "We would like to invite you to get to know us") without mentioning a specific date.
Examples:

- "Are you available for a viewing this Friday at 1pm?"
- "Why don't you give me a call tomorrow morning before 2pm and we can arrange a viewing."
- "Thank you for your application! If you like, we can arrange a viewing."
- "Contact ... and you can arrange to meet to get to know each other."
- "Please call me on ... so we can get to know each other."

Other Responses: Responses where the advertiser or potential roommate requests additional information about the applicant, indicating a potential interest in the applicant, are classified as "other" responses.
Examples:

- "Nice to hear from you. When do you plan to move in and for how long?"
- "Sounds good ;) How long is your master's program?"
- "The room is more of a functional [non-communal] flat share, as there are no common rooms. Are you still interested?"
- "How will you finance your studies? The room is not furnished."
- "Before I make an appointment to view the room, I would like you to answer some important questions: When would you like to start renting? Are you already in ... or where do you currently live? Have you already checked the travel time from the shared apartment to the university?"
- "Are you a vegetarian?"
- "Do you have an email address where I can send you more information?"
- "Sorry, the room is already taken. But there will be another room available in November [...]. If this is something for you, please contact me [...]."
- "Thank you for your application! Would you be willing to move in by 09/01?"

B.4 Study II: Application Text

The following is a loose translation of the application text. The original text was sent out in German. The italicized terms represent variables that are randomly added to the text based on the randomly selected ethnicity and information conditions.

> Hello *advertiser*,
> I just saw the ad for the vacant room in the shared apartment and I'm very interested. My name is *full_name*. I'm 24 and I've just started a Master's in Business Administration, which is why I'm now looking for a room. I've been subletting so far, but I'm now looking for something longer-term in *city*.
> I thought I'd give you a quick overview of who I am: I don't smoke, I've already lived in a shared apartment, and I know what a cleaning schedule is. In my spare time, I enjoy meeting up with friends for coffee or a drink, going for a jog, or travel the world. But of course, watching a good TV series from time to time is also a must.
> I would be happy to introduce myself and see the room. I am flexible with dates.
> Best regards
> *first_name*
> P.S.: If you want to get a picture of me, here's a link to my Instagram profile: https://www.instagram.com/*instagram_profile*

The hyperlink for the social media profile is automatically converted by the platform into a clickable link, facilitating the accessibility of additional social media information at minimal cost. Upon clicking the link on a mobile device with the Instagram mobile application installed and the user logged in, the profile is typically opened directly within the Instagram mobile application where all profile information is directly accessible.

On a computer, the profile opens in a new browser window/tab, allowing the user to access the entire profile (including the username, name, biographic information, and the first image of every post[1]) without being logged in. However, additional information, such as a list of subscribers and subscriptions, as well as additional images within a post or image descriptions, are only accessible when logged in.

[1]See screenshots of the entire profiles in Figures B.8a, B.8b, B.9a, and B.9b.

B.5 Screenshots

B.5.1 Social Media Profiles

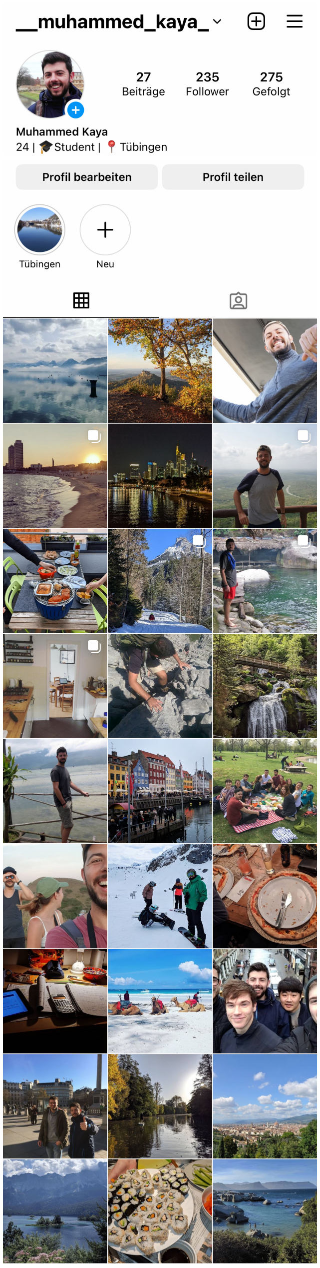


Figure B.8a: Minority SM Profile Without Visual Minority Stereotypes

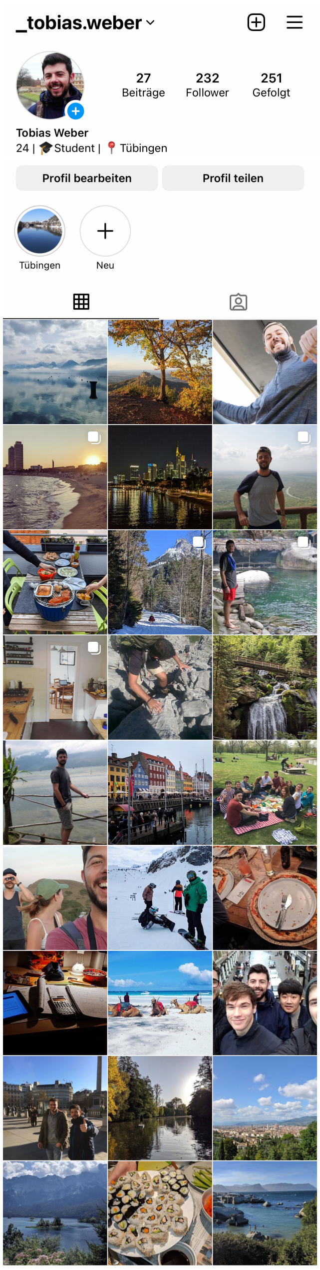


Figure B.8b: Majority SM Profile Without Visual Minority Stereotypes

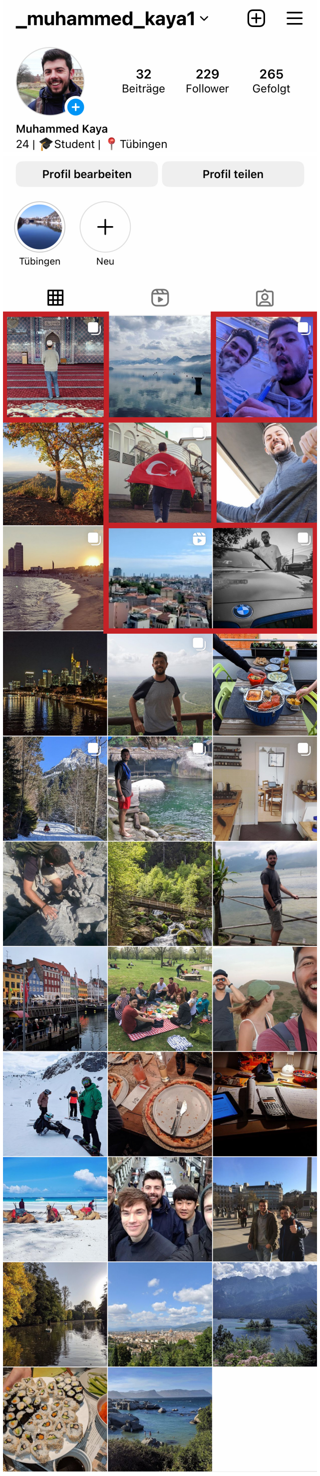


Figure B.9a: Minority SM Profile With Visual Minority Stereotypes (red bordered box)

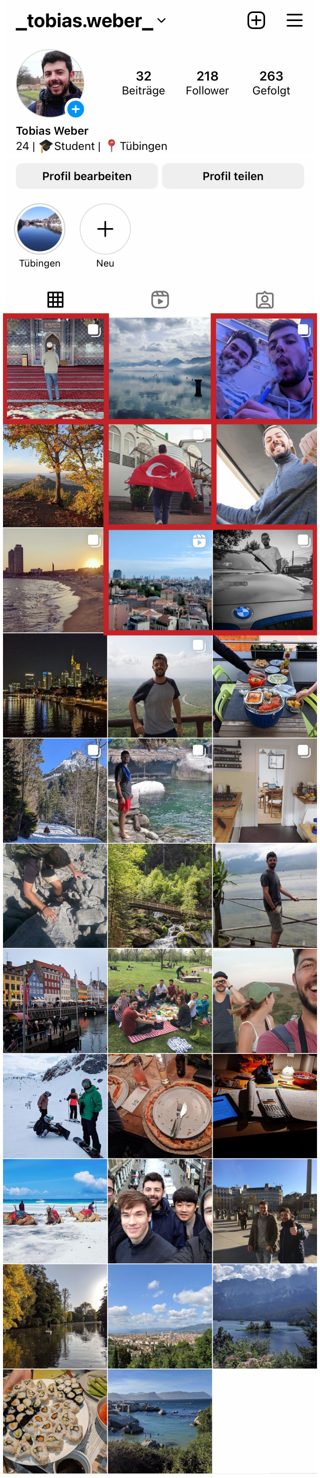


Figure B.9b: Majority SM Profile With Visual Minority Stereotypes (red bordered box)

B.5.2 Visual Minority Stereotypes on Social Media

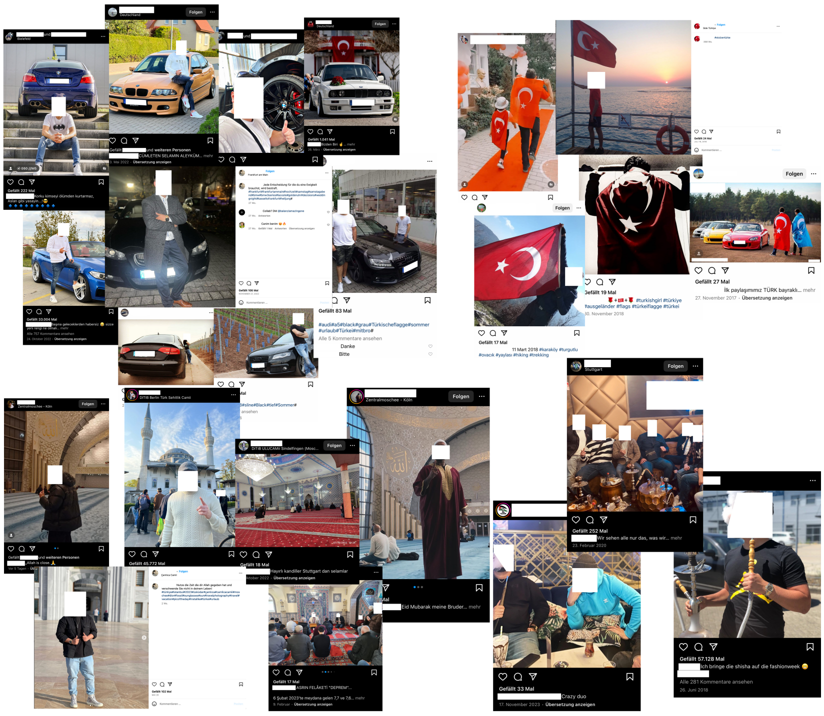


Figure B.10: Screenshots of Similar Posts

B.5.3 Friend Suggestions & Treatment Notifications

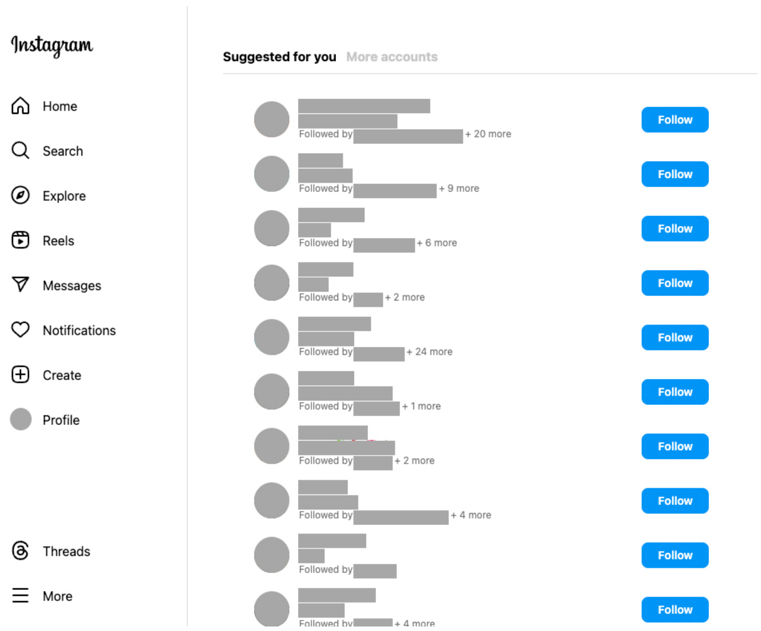


Figure B.11: Screenshot of Friend Suggestions (Desktop App)

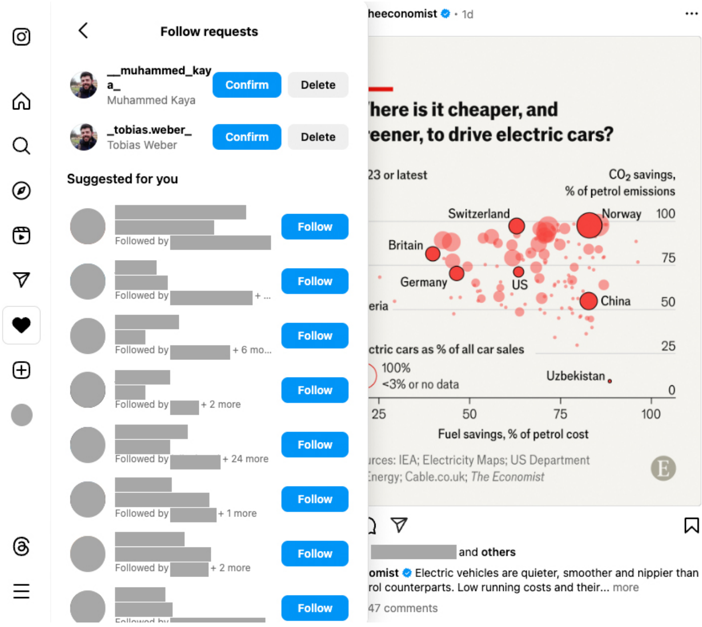


Figure B.12a: Screenshot of (Example) Treatment Notification (Desktop App)

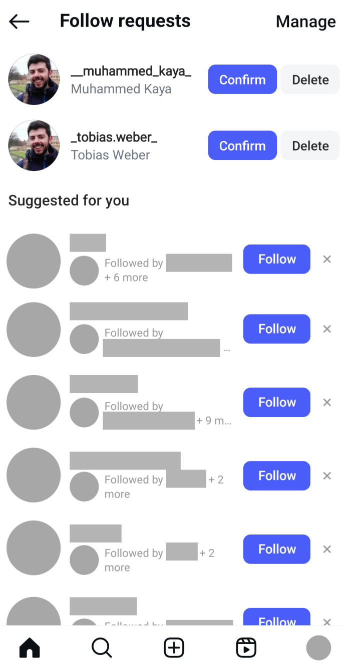


Figure B.12b: Screenshot of (Example) Treatment Notification (Mobile App)

B.6 Online Pilot Experiments

Prior to the start of the experiments, we conducted two randomized online pilot experiments to assess the effectiveness of the treatment conditions in manipulating relevant beliefs. In addition, our objective was to validate our experimental design, especially with regard to (visual) minority stereotypes and the design of the social media (SM) profiles on the photo-sharing platform Instagram.

Firstly, we tested whether the male names selected from Chapter 2 accurately represented the intended ethnic minority or majority status in the absence of manipulated social media profiles. Secondly, we wanted to find out what characteristics one would typically associate with a person having an ethnic minority (majority) name (open-ended questions), for a population similar to the later subject pool in the field experiment.

Using this information and previous research (Ossenberg, 2019), we created visual minority stereotypes representing the most frequently mentioned characteristics from the first online pilot experiment. In a second online pilot experiment, we tested whether these visual minority stereotypes posted on social media profiles was effective in manipulating relevant beliefs.

Subsequently, we conducted an additional survey asking participants to subscribe to our profiles and rate their perceived realism.

B.6.1 First Online Pilot Experiment

The objective of the first online pilot experiment was to validate the attribution of treatment names from Chapter 2 to their intended ethnic origin, religious affiliation, and gender. Additionally, we aimed to investigate the characteristics commonly associated with individuals of Turkish origin or Turkish migration background and German origin in general and in particular for a given name – for a population resembling the later population in the experiment. Furthermore, participants were asked to evaluate predefined characteristics and the potential compatibility as a roommate without any additional (social media) information.

Survey Design In the *first stage*, we randomly assigned participants to either the Turkish-sounding male name ("Muhammed Kaya") or German-sounding male name ("Tobias Weber"). Then, we asked subjects to assess the ethnic origin of a person with the given name using open-ended questions.[2] Additionally, we asked participants to rate the name in terms of gender, ethnicity, and religious affiliation.[3]

In the *second stage*, we randomly assigned participants to a more general ethnic minority (ethnic majority) condition and asked them to think of characteristics that one would generally attribute to a person of Turkish (German) origin unrelated to a specific name.

In the *third stage*, as in the beginning, participants were again assigned to one of the two specific names of the experiment (see above). To control for possible inter-rater reliability, a participant who was previously assigned to the minority is now asked to rate the name of the majority and vice versa. Subsequently, we asked participants to evaluate the extent to which a given ethnic minority (majority) name adheres to the ethnic minority (majority) characteristics previously mentioned in an open-ended question. Then, we asked to rate the likelihood of the person with the given name possessing certain traits, such as punctuality, patriotism, religiosity, and other common Turkish stereotypes held by Germans (Ossenberg, 2019) on a 7-point Likert scale. In addition, we asked participants how likely they would be to invite a person with the given name to a viewing if they had a room available in their shared apartment.

[2]The first stage of this online pilot was also described in Chapter 2 (see Section A.5.2, p. 168) where additional treatment conditions were included for a broader set of names as used in Chapter 2.

[3]Respondents were asked to rate on a Likert scale ranging from 1 (very unlikely) to 7 (very likely).

Finally, we asked participants to provide us with demographic information including gender, age, occupation, household size, experience living in shared apartments, nationality, and whether they have a migration background, i.e. if (one of) their parents emigrated to Germany.

Results The pilot experiment was conducted online in June 2023 using a professional survey platform. We recruited a total of $n = 1,725$ participants through a university-wide circular mail. We excluded 324 observations from our analysis since these did not pass certain quality tests.[4] Although the sample may not be nationally representative, it is substantially similar to the later subject pool of the field experiment in terms of demographic characteristics. Roommates are typically young students, who also make up the majority of the recipients of the circular mail. Specifically, 67.4% of participants indicated that they are students, while 28.6% indicated that they are already (part-time) employed, mainly belonging to faculty members and administrative staff.

Furthermore, the survey respondent's median age is 24 years, while the average age is 28.8 years (see Table B.29 for summary statistics). The majority of respondents (73.5%) identified as female. Additionally, most participants (66.4%) reported living in a shared apartment currently or in the past. 92.4% of respondents indicated a German nationality. Turkish was the second most commonly mentioned nationality, but only to a small extent (1.3%). However, 25.8% of the respondents reported having a migration background, meaning that at least one of their parents was born abroad. Overall, 2.4% of respondents reported having either Turkish citizenship or a Turkish migration background.

The results of the *first stage* of the first online pilot experiment are presented in Table B.30. Our data suggests that the randomized treatment successfully manipulates relevant beliefs, resulting in the Turkish name being primarily categorized as Turkish in both the open-ended and closed questions using a 7-point Likert scale. The same can be observed for the German name (see column 4 of Table B.30). Although the Turkish name may also suggest an Arabic or Middle Eastern origin, it is important to note that it has the largest mean values in all cases. This is in line with the results of different OLS regressions of the manually categorized open-ended questions on the origin of the treatment name on a dummy that equals one if the participant was randomly selected to rate the Turkish name and selected interactions (see Table B.31).

The results indicate that respondents associate the Turkish name with a Turkish origin as intended (0.443, $p = 0.000$; see column 1 of Table B.31). The treatment effects on other origins, such as Arabic or Middle Eastern (see columns 2 and 3 of Table B.31) are lower in magnitude. All estimations include demographic control variables, such as age, occupation, or nationality/migration background, and additional survey controls, such as the time that respondents needed in order to complete the experiment. Standard errors are clustered on the individual level.

Furthermore, both names are characterized as male names (see Table B.30). Additionally, the Turkish (German) name is strongly associated with being a Muslim (Christian) name. Simple two-sample t-tests reveal statistically significant differences in the means of the characteristics of Turkish and German names. Overall, it is highly unlikely that participants in the field experiment would fail to associate the intended ethnicity and/or gender.

The results of the *second and third stage* of the online pilot experiment are reported in Table B.32. It shows the characteristics most frequently mentioned by subjects when asked to enter (open-ended question) characteristics associated with someone with a Turkish migration background (stage two) or the Turkish male name used in the later experiment (stage three). We manually categorized the open-ended questions using a content analysis, following a categorization scheme developed from the respondent's answers (Carley, 1993; Jackson & Trochim,

[4] We dropped observations that completed the questionnaire too fast, i.e., whose relative speed index were too high (Leiner, 2019). Additionally, we only included responses from German-speaking participants, as the questionnaire was designed to capture ethnic minority stereotypes held by the ethnic majority in Germany. Observations where participants indicated that they did not complete the survey truthfully or thoroughly were excluded. We also excluded subjects who entered nonsense or used abusive language in open-ended questions.

2002). The results of the present study are presented in columns 1 and 3 of Table B.32, while columns 2 and 4 display the results of a survey conducted by Ossenberg (2019) between 2014 and 2016 in Germany. The visual minority stereotypes were designed based on the results of both studies (see Section 3.3.1.1).

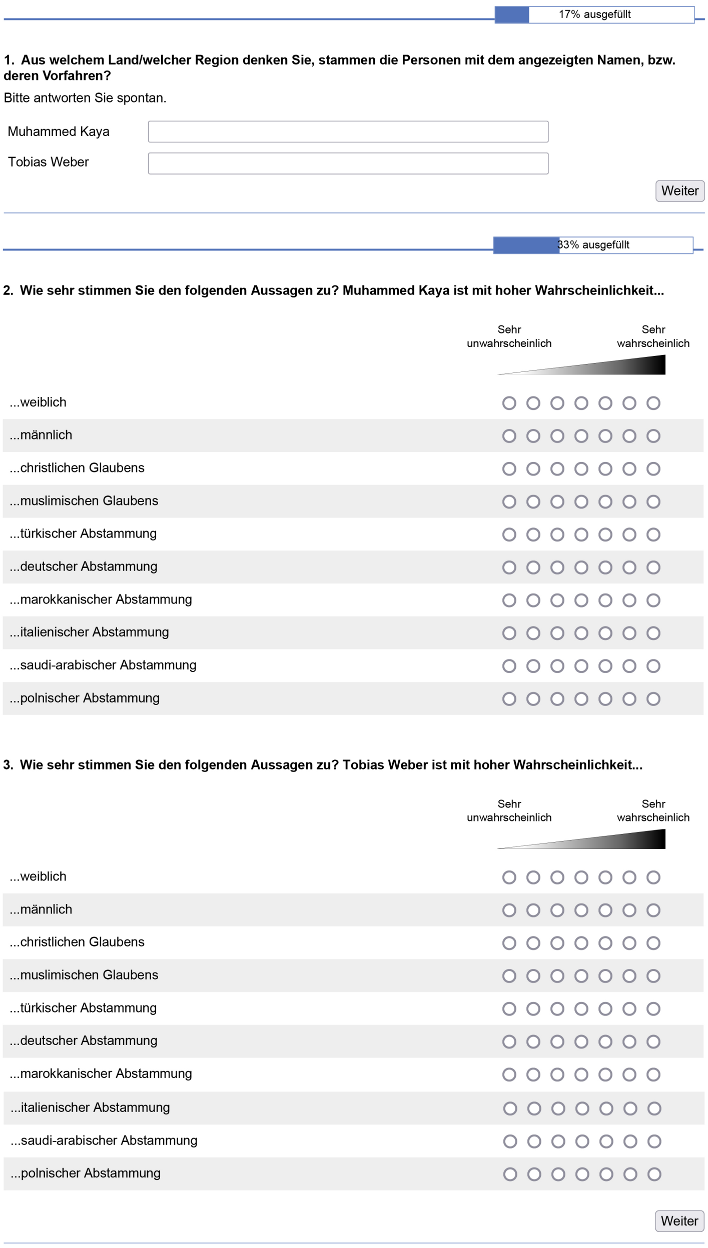
As additional controls, participants were asked to evaluate the Turkish (German) name based on predefined characteristics, such as punctuality or patriotism, based on Ossenberg (2019) using a 7-point Likert scale.[5] Consistent with prior research, a person with the Turkish-sounding name is generally perceived as more patriotic, religious, and tradition-bound compared to a person with the German-sounding name (see Table B.33). The latter is more likely to display punctuality, discipline, and environmental conscientiousness. The differences in means are statistically significant (see column 7 of Table B.33).

Moreover, the Turkish-sounding name has a slightly higher likelihood of being invited to view a room in a shared apartment and meet the potential roommates. This finding contradicts nearly all correspondence tests on discrimination (Bertrand & Duflo, 2017; Lippens et al., 2023). This may be because participants are aware that their behavior is being studied.

Of course, the deception of participants in experimental studies should be avoided whenever possible. However, comparisons between the survey data and experimental results indicate that education on sensitive topics, such as discrimination, can elicit socially desirable behavior, thereby compromising the validity of the findings.

This fact highlights the importance of conducting experimental research without informing participants that they are part of a study. However, it is important to consider the lack of voluntary participation and informed consent when designing experiments due to the substantial ethical implications (see Section B.8 for ethical considerations of this research, as well as Zschirnt (2019)).

[5] The questionnaire was designed so that participants could not modify their open-ended responses once they had been asked closed-ended questions, which may have covered similar topics as the open-ended questions.



17% ausgefüllt

1. Aus welchem Land/welcher Region denken Sie, stammen die Personen mit dem angezeigten Namen, bzw. deren Vorfahren?

Bitte antworten Sie spontan.

Muhammed Kaya

Tobias Weber

Weiter

33% ausgefüllt

2. Wie sehr stimmen Sie den folgenden Aussagen zu? Muhammed Kaya ist mit hoher Wahrscheinlichkeit...

	Sehr unwahrscheinlich						Sehr wahrscheinlich
...weiblich	○	○	○	○	○	○	○
...männlich	○	○	○	○	○	○	○
...christlichen Glaubens	○	○	○	○	○	○	○
...muslimischen Glaubens	○	○	○	○	○	○	○
...türkischer Abstammung	○	○	○	○	○	○	○
...deutscher Abstammung	○	○	○	○	○	○	○
...marokkanischer Abstammung	○	○	○	○	○	○	○
...italienischer Abstammung	○	○	○	○	○	○	○
...saudi-arabischer Abstammung	○	○	○	○	○	○	○
...polnischer Abstammung	○	○	○	○	○	○	○

3. Wie sehr stimmen Sie den folgenden Aussagen zu? Tobias Weber ist mit hoher Wahrscheinlichkeit...

	Sehr unwahrscheinlich						Sehr wahrscheinlich
...weiblich	○	○	○	○	○	○	○
...männlich	○	○	○	○	○	○	○
...christlichen Glaubens	○	○	○	○	○	○	○
...muslimischen Glaubens	○	○	○	○	○	○	○
...türkischer Abstammung	○	○	○	○	○	○	○
...deutscher Abstammung	○	○	○	○	○	○	○
...marokkanischer Abstammung	○	○	○	○	○	○	○
...italienischer Abstammung	○	○	○	○	○	○	○
...saudi-arabischer Abstammung	○	○	○	○	○	○	○
...polnischer Abstammung	○	○	○	○	○	○	○

Weiter

Figure B.13: Screenshot of the First Stage of the First Online Pilot

Table B.29: First Online Pilot Experiment: Summary Statistics

	Mean	S.D.	Min	Max	Obs.
Treatment 1: Turkish name first	0.481	0.499	0	1	1,401
Treatment 2: Turkish ethnicity	0.498	0.500	0	1	1,401
Treatment 3: Turkish name	0.502	0.500	0	1	1,401
Female	0.735	0.441	0	1	1,262
Age	28.77	11.57	16	79	1,200
Student	0.674	0.469	0	1	1,265
Employed (full-time)	0.169	0.375	0	1	1,265
Employed (part-time)	0.117	0.322	0	1	1,265
Shared apartment experience	0.664	0.473	0	1	1,255
Shared apartment experience (in years)	3.833	3.160	1	25	797
Number of roommates	3.759	4.638	1	53	1,214
German	0.924	0.265	0	1	1,201
Turkish	0.0236	0.152	0	1	1,401
Migration background	0.258	0.438	0	1	1,223
Time (in seconds)	423.5	171.8	105	1,168	1,401
Finished	0.903	0.296	0	1	1,401
Share of missing responses	2.525	2.559	0	35	1,401

Note: The table reports summary statistics for the first online pilot experiment.

Table B.30: First Online Pilot Experiment: Summary Statistics (First Stage)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Mean	S.D.	Obs.	Mean	S.D.	Obs.	Diff. in Means (p-value)
	Turkish name			German name			
German origin	0.0452	0.208	332	0.972	0.165	357	-0.93*** (0.000)
Turkish origin	0.461	0.499	332	0	0	357	0.46*** (0.000)
Arabic origin	0.274	0.447	332	0	0	357	0.27*** (0.000)
Middle Eastern origin	0.111	0.315	332	0	0	357	0.11*** (0.000)
Other origin	0.108	0.311	332	0.0280	0.165	357	0.08*** (0.000)
Female	1.359	0.854	334	1.154	0.487	357	0.21*** (0.000)
Male	6.749	0.704	334	6.868	0.476	357	-0.12*** (0.009)
Christian	2.569	1.160	334	4.905	1.113	357	-2.34*** (0.000)
Muslim	5.620	1.075	334	2.224	1.065	357	3.39*** (0.000)
Turkish	5.326	1.308	334	1.896	1.035	357	3.43*** (0.000)
German	2.886	1.390	334	6.216	1.058	357	-3.33*** (0.000)
Moroccan	4.461	1.543	334	1.835	0.962	357	2.63*** (0.000)
Italian	2.186	1.060	334	2.375	1.330	357	-0.19** (0.039)
Saudi Arabian	4.850	1.477	333	1.765	0.983	357	3.09*** (0.000)
Polish	1.916	1.048	334	2.936	1.515	357	-1.02*** (0.000)

Note: The table reports summary statistics for the first stage of the first online pilot experiment. The first five variables are dummy variables that equal 1 if the respondent indicated an origin that matches the respective category. The remaining variables are continuous variables measured on a 7-point Likert scale ranging from very unlikely to very likely. Column 7 displays a pairwise comparison of differences in means using two-sample t-tests with equal variances. P-values are given in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.31: First Online Pilot Experiment: Treatment Name Manipulation Checks (First Stage)

	(1) Name origin: Turkey	(2) Name origin: Arabic	(3) Name origin: Middle East	(4) Name origin: Turkey
Treatment: Turkish name	0.443***	0.276***	0.125***	0.186
	(0.0308)	(0.0280)	(0.0208)	(0.114)
Female	0.0423	-0.00389	-0.0166	0.0450
	(0.0332)	(0.0306)	(0.0242)	(0.0331)
Age	-0.000658	-0.00146	0.00258*	-0.000840
	(0.00185)	(0.00173)	(0.00142)	(0.00187)
Student	0.0923	-0.0436	0.0298	0.0843
	(0.0581)	(0.0553)	(0.0347)	(0.0588)
Employed	0.00306	0.0376	0.00371	0.00836
	(0.0516)	(0.0527)	(0.0380)	(0.0521)
Shared apartment	-0.00574	-0.0247	0.00522	-0.0102
	(0.0322)	(0.0300)	(0.0209)	(0.0322)
Roommates	-0.00318	0.00301	0.00133	-0.00274
	(0.00231)	(0.00305)	(0.00199)	(0.00236)
German	0.127**	-0.158***	0.0180	0.00328
	(0.0579)	(0.0600)	(0.0361)	(0.0219)
Turkish	0.240**	-0.207***	-0.0466	-0.0292
	(0.117)	(0.0578)	(0.0340)	(0.0269)
Migration background	0.0406	-0.0464	-0.0181	-0.00619
	(0.0360)	(0.0293)	(0.0207)	(0.00980)
Treatment × German				0.239**
				(0.111)
Treatment × Migr. backgr.				0.112
				(0.0800)
Treatment × Turkish				0.429**
				(0.193)
Constant	-0.0259	0.242	-0.201**	0.123
	(0.170)	(0.161)	(0.0908)	(0.162)
Additional Survey Controls	Yes	Yes	Yes	Yes
Observations	546	546	546	546
R-squared	0.331	0.193	0.092	0.341

Note: The table reports OLS regressions using different dependent variables derived from an open-ended question on the perceived origin of a given name. Responses are coded as dummy variables that equal 1 if the respondent indicated an origin which corresponds to the respective category. Additional survey controls include the time required for completion, a dummy variable indicating whether the survey was fully completed, and an indicator for the first randomly selected treatment condition. Standard errors are clustered at the individual level. Robust standard errors in parenthesis. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table B.32: First Online Pilot Experiment: Stereotypes (Second & Third Stage)

(1) Turkish characteristics (this study)	%	(2) Germans about Turks (Ossenberg, 2019)	%	(3) German characteristics (this study)	%	(4) Germans about Germans (Ossenberg, 2019)	%
agreeable	10.95	religious	54.62	punctual	14.37	punctual	52.44
family-oriented	10.03	family-oriented	53.06	organized	7.33	bureaucratic	38.47
religious	9.70	tradition-bound	49.62	accurate	7.20	fond of drinking	36.56
darker skin tone	8.35	patriotic	47.04	conscientious	6.97	disciplined	35.87
extravert	7.96	hospitable	46.40	white	6.24	conscientious	35.49
open	5.39	national pride	45.97	narrow-minded	5.77	orderly	33.81
hospitable	4.70	proud	40.22	hardworking	5.58	hardworking	32.67
tradition-bound	4.01	sense of belonging	31.61	german	5.58	reliable	28.39
hardworking	3.45	sociable	31.24	agreeable	4.58	good organizers	28.32
loud	3.22	impulsive	25.91	negative	3.52	thorough	28.09

Note: The table reports the most frequently mentioned characteristics of the experiment participants regarding common Turkish and German traits, as well as traits associated with individuals who have the name of our fictional applicants. Columns 1 and 3 show the results of the current study (second and third stage pooled) together with the frequencies of these traits, while columns 2 and 4 display the results of a study by Ossenberg (2019).

Table B.33: First Online Pilot Experiment: Summary Statistics (Third Stage)

	(1) Mean	(2) S.D.	(3) Obs.	(4) Mean	(5) S.D.	(6) Obs.	(7) Diff. in Means (p-value)
	Turkish name			German name			
Punctual	4.154	1.099	656	4.966	1.264	640	-0.81*** (0.000)
Patriotic	4.703	1.274	656	3.486	1.293	640	1.22*** (0.000)
Religious	5.200	1.082	656	3.472	1.127	640	1.73*** (0.000)
Disciplined	4.491	1.087	656	4.745	1.175	640	-0.25*** (0.000)
Tradition-bound	5.308	1.155	656	3.812	1.298	639	1.49*** (0.000)
Environmentally conscious	3.736	1.113	656	4.314	1.292	640	-0.58*** (0.000)
Callback	4.229	0.970	323	4.072	0.862	335	0.16** (0.028)

Note: The table reports summary statistics for the third stage of the first online pilot experiment. The first six variables are measured on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). Callback is measured on a 5-point Likert scale ranging from 1 (definitely not) to 5 (definitely). Column 7 displays a pairwise comparison of differences in means using two-sample t-tests with equal variances. P-values are given in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.6.2 Second Online Pilot Experiment

The purpose of the second online pilot experiment was to validate whether the treatment conditions successfully manipulated relevant beliefs regarding perceptions of visual minority stereotypes and ethnicity. Participants were randomly assigned to the manipulated social media information and asked to assess the applicant and personal characteristics that can be inferred from the visual information in open- and closed-ended questions. Additionally, participants were asked to evaluate the applicant's attractiveness, fit as a potential roommate, and the authenticity of the social media profiles.

Survey Design We randomly assigned participants to one of the four conditions: Turkish vs. German name, with vs. without Turkish stereotypical images using a between-subjects design. Participants were then presented with the entire social media profile of the respective condition and asked to examine it in detail. Afterwards, the most recently posted images from the respective profile were presented and participants were asked to closely examine all of the images. This included all stereotypical images in the conditions where Turkish stereotypical images were presented.[6]

The full profiles and respective images were presented to the subjects in the exact same way as they would access it if they received a link in an application, as in the later experiment. Thus, subjects had access to the same level of information as in the later experiment, including names, biographical information, subscribers, captions, hashtags, etc. (see Sections 3.3.2 and 3.3.3).

Participants were asked to evaluate the profile owner's physical appearance on a 5-point Likert scale, ranging from very attractive to very unattractive. Following this, participants were instructed to provide spontaneous descriptions of potential characteristics that describe the individual depicted in the social media profile (open-ended questions). As an additional manipulation check, participants were asked to evaluate the person's religious affiliation and ethnic background, as well as the likelihood that the individual matched pre-defined characteristics such as punctuality and patriotism (Ossenberg, 2019).[7]

Additionally, we asked participants how likely they would be to invite the individual from the social media profile to view the spare room and attend a get-to-know-you meeting with potential roommates (*callback*) if they hypothetically lived in a shared apartment with a vacant room. Subsequently, subjects were instructed to rate the realism and authenticity of the profile.[8] Throughout the online pilot, participants were intentionally able to view the social media profile or images again.

Finally, participants were asked to provide demographic information, including gender, age, experience living in shared apartments, household size, nationality, and frequency of use of common social media platforms.

Results The pilot experiment was conducted online between December 2022 and January 2023 using a professional survey platform. We recruited a total of $n = 506$ students from our university. We excluded 36 observations from our analysis because they only examined social media profiles without providing any information or completing the survey.

On average, 57.6% of respondents identified as female with an average age of 25.6 years (see Table B.34). More than half of the respondents lived at the time of the online pilot in a shared apartment since an average of 3.2 years. 85.7% (1.9%) of respondents indicated a German (Turkish) nationality. Participants most frequently use the social media platform Instagram, on which we created the fictitious profiles, followed by Snapchat and Tiktok. Participants reported using Instagram several times a day (median).

Additional descriptive results show that manipulating relevant beliefs through social media information resulted in different outcomes depending on the treatment condition. Table B.35 shows that subjects attributed different characteristics to the same applicant based on the ethnicity signal and ethnic minority stereotypes.

In order to analyze the open-ended responses, we manually categorized the entries using a content analysis, following a categorization scheme developed from the subject's answers (Carley, 1993; Jackson & Trochim, 2002). The results indicate that participants attributed a higher number of Turkish characteristics to a profile owner with a Turkish-sounding name (see first part of Table B.35). The same is true for characteristics indicating a Turkish origin, with

[6]For technical reasons and to improve accessibility, we did not include the short video ("Instagram Reel") posted on the profile. The video is a nine-second clip showing the city of Istanbul with the call for prayer audible in the background.

[7]Respondents were asked to rate on a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

[8]Participants were asked to rate the callback item on a Likert scale from 1 (No, definitely not) to 5 (Yes, definitely). The realism item was rated on a scale from 1 (Not realistic at all) to 5 (Extremely realistic) without a neutral option.

the exception for the profile with a German name but with Turkish stereotypical images, where a Turkish origin is indicated by the respective pictures posted on the profile.

Furthermore, participants mentioned religious or patriotic characteristics more frequently when assigned to profile conditions that display ethnic minority stereotypes through religious or patriotic images (see columns 5-8 of Table B.35). The same is true for pre-defined characteristics, such as being Christian or Muslim, patriotic, religious, or tradition-bound (see second part of Table B.35). Comparatively lower differences are reported for stereotypes commonly associated with Germans, such as punctuality and discipline (see Table B.32 from the first online pilot).

Although each profile displays the same individual, participants attribute lower average attractiveness scores to profiles that display ethnic minority stereotypes. Profiles with a Turkish name and without Turkish stereotypical images have the highest average callback rate, followed by a profile with a German name in the same condition. These findings are in line with the ones from Chapter 2 of this thesis. The lowest average callback rate is reported for the condition where a profile with a Turkish name and Turkish stereotypical images was shown to subjects.

Generally, all profiles are perceived to be rather realistic and the perceived realism does not significantly vary between treatment conditions (see third part of Table B.35). Note that the 5-point Likert scale for the realism item deliberately excluded a neutral option. As a result, the vast majority (89.9%) of participants rated the profiles as somewhat, very, or extremely realistic. Only 9.9% of subjects rated the profiles as conditionally realistic, while only one respondent (0.2%) rated the respective profile as not realistic at all.

In Table B.36, we perform multiple Wilcoxon rank-sum tests to determine statistically significant differences in means based on treatment conditions. The results indicate a statistically significant difference in the perception of individuals regarding a Turkish origin or being Muslim when comparing the ethnicity conditions (see Panel A of Table B.36) and the stereotype conditions (see Panel B of Table B.36). Panel C and D of Table B.36 report means and differences of stereotype conditions by ethnicity and vice versa.

Overall, the descriptive results suggest that participants attributed Turkish and Muslim characteristics significantly more often to profiles with a Turkish-sounding name and to profiles showing Turkish stereotypical images – indicating that we successfully manipulated the relevant information representing ethnicity and ethnic minority stereotypes.

Table B.37 presents the results of multiple OLS regressions of different treatment and control variables on the pre-defined characteristics of being perceived as Turkish or Muslim (measured on a 5-point Likert scale). All models include various demographic control variables, such as gender, age, and nationality, as well as additional survey controls, such as the time required to complete the online pilot.

The results are broadly in line with the descriptive results, indicating that a Turkish name (*Treatment: Turkish*) has a statistically significant positive effect on the perception of Turkish origin or being Muslim (1.144, $p = 0.000$ and 2.192, $p = 0.000$; see columns 2 and 4 of Table B.37). Similarly, stereotypical images (*Treatment: Stereotypes*) have on average a statistically significant positive effect on both perception variables, which is consistent across all models.

The negative coefficients of the treatment interaction terms *Treatment: Turkish* $\times$ *Stereotypes* indicate that the combined effect of the Turkish name treatment and stereotypical images is less than the sum of their individual effects on both perception variables. This suggests that the presence of stereotypes may counteract or moderate the positive impact of the Turkish treatment to some extent. Thus, coupling stereotypical images with Turkish names may attenuate the treatment effect on the perception of a Turkish origin.

Regarding the likelihood of a callback, Table B.38 presents the results of regressing treatment and control variables on the callback probability. Column 1 presents the results without interaction effects, column 2 with the treatment interaction effect, and column 3 with all interaction effects, including interactions of the respondent's nationality and the treatment variables.

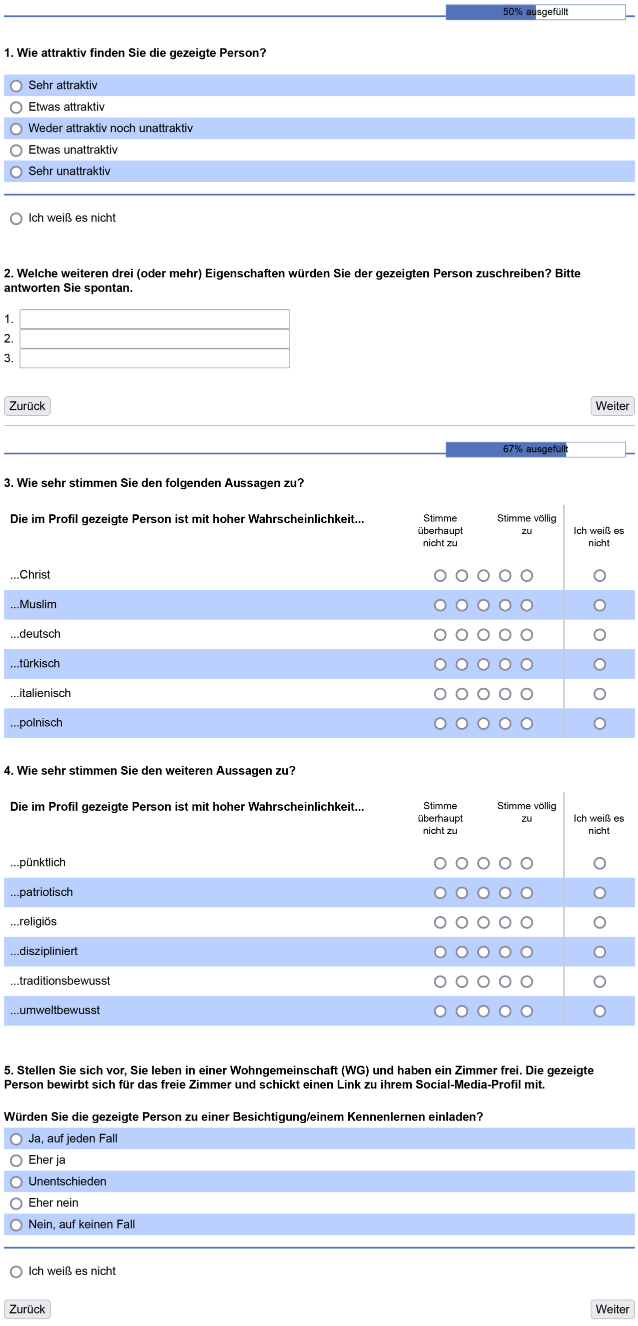
The data suggest that the likelihood of receiving a callback is primarily driven by the treatment conditions where the subject's were presented a profile with images representing ethnic minority stereotypes. The coefficients for the Turkish treatment are not statistically significant and are inconsistent between the models, varying greatly in both magnitude and direction of effect.

The presence of stereotypical images (*Treatment: Stereotypes*) has a statistically significant negative effect on the average likelihood of receiving a callback. However, this effect is not consistent across all models. The interaction effect between both treatment variables (*Treatment: Turkish* $\times$ *Stereotypes*) appears to be more robust. The coefficients of -0.389 ($p = 0.049$) to -0.432 ($p = 0.029$) indicate that the combined effect of having a Turkish name and stereotypical images is less negative than the sum of their individual effects on the likelihood of receiving a callback (see columns 2 and 3 of Table B.38). Therefore, it appears that stereotypical images are the primary factor leading to lower callback rates.

Moreover, the presence of stereotypical images in profiles with Turkish names may partially counteract the potential negative bias associated with the Turkish name. Thus, stereotypical images act as moderators and do not exacerbate the negative impact, but rather mitigate it. Therefore, the stereotypical images may challenge or modify pre-existing biases associated with ethnic names, emphasizing the significance of context and individual interpretation in shaping perceptions.

In addition, the coefficient of the interaction term *Treatment: Stereotypes* $\times$ *Turkish* is statistically significant and positive (1.468, $p = 0.077$; see column 3 of Table B.38). This suggests that, despite the generally negative effect of stereotypes on callback probability, the likelihood of receiving a callback for respondents with Turkish nationality may be increased. A reason could again be subject's socially desirable behavior as discussed briefly in Section B.6.1.

In addition, a greater level of perceived physical attractiveness is, on average, associated with a higher probability of receiving a callback. This effect is statistically significant in all models.



50% ausgefüllt

1. Wie attraktiv finden Sie die gezeigte Person?

- ○ Sehr attraktiv
- ○ Etwas attraktiv
- ○ Weder attraktiv noch unattraktiv
- ○ Etwas unattraktiv
- ○ Sehr unattraktiv

- ○ Ich weiß es nicht

2. Welche weiteren drei (oder mehr) Eigenschaften würden Sie der gezeigten Person zuschreiben? Bitte antworten Sie spontan.

1.
2.
3.

Zurück | Weiter

67% ausgefüllt

3. Wie sehr stimmen Sie den folgenden Aussagen zu?

Die im Profil gezeigte Person ist mit hoher Wahrscheinlichkeit...	Stimme überhaupt nicht zu				Stimme völlig zu	Ich weiß es nicht
...Christ	○	○	○	○	○	○
...Muslim	○	○	○	○	○	○
...deutsch	○	○	○	○	○	○
...türkisch	○	○	○	○	○	○
...italienisch	○	○	○	○	○	○
...polnisch	○	○	○	○	○	○

4. Wie sehr stimmen Sie den weiteren Aussagen zu?

Die im Profil gezeigte Person ist mit hoher Wahrscheinlichkeit...	Stimme überhaupt nicht zu				Stimme völlig zu	Ich weiß es nicht
...pünktlich	○	○	○	○	○	○
...patriotisch	○	○	○	○	○	○
...religiös	○	○	○	○	○	○
...diszipliniert	○	○	○	○	○	○
...traditionsbewusst	○	○	○	○	○	○
...umweltbewusst	○	○	○	○	○	○

5. Stellen Sie sich vor, Sie leben in einer Wohngemeinschaft (WG) und haben ein Zimmer frei. Die gezeigte Person bewirbt sich für das freie Zimmer und schickt einen Link zu ihrem Social-Media-Profil mit.

Würden Sie die gezeigte Person zu einer Besichtigung/einem Kennenlernen einladen?

- ○ Ja, auf jeden Fall
- ○ Eher ja
- ○ Unentschieden
- ○ Eher nein
- ○ Nein, auf keinen Fall

- ○ Ich weiß es nicht

Zurück | Weiter

Figure B.14: Screenshot of the Second Online Pilot

Table B.34: Second Online Pilot Experiment: Summary Statistics

	Mean	S.D.	Min	Max	Obs
Treatment: Turkish with Minority Stereotypes	0.270	0.445	0	1	470
Treatment: German with Minority Stereotypes	0.238	0.426	0	1	470
Treatment: Turkish without Minority Stereotypes	0.249	0.433	0	1	470
Treatment: German without Minority Stereotypes	0.243	0.429	0	1	470
Female	0.576	0.495	0	1	462
Age	25.55	9.975	15	97	447
Shared apartment experience	0.527	0.500	0	1	463
Shared apartment experience (in years)	3.199	2.689	0	19	236
German	0.857	0.350	0	1	470
Turkish	0.019	0.137	0	1	470
Foreign	0.107	0.310	0	1	439
Double nationality	0.0234	0.151	0	1	470
Facebook usage	2.043	1.539	1	6	465
Instagram usage	4.664	1.847	1	6	464
Twitter usage	1.614	1.237	1	6	464
Snapchat usage	3.269	2.125	1	6	465
Tiktok usage	2.404	1.989	1	6	465
Participated in Social Media study	0.0917	0.289	0	1	458
Time (in seconds)	344.2	150.4	121	1,067	470
Finished	0.985	0.121	0	1	470
Share of missing responses	6.777	2.627	0	21	470

Note: The table reports summary statistics for the second online pilot experiment.

Table B.35: Second Online Pilot Experiment: Summary Statistics on Characteristics

	(1) Mean	(2) Obs.	(3) Mean	(4) Obs.	(5) Mean	(6) Obs.	(7) Mean	(8) Obs.
	Turkish name without Minority Stereotypes		German name without Minority Stereotypes		Turkish name with Minority Stereotypes		German name with Minority Stereotypes	
Categorization of characteristics (open-ended question)								
Turkish characteristics	0.0598	117	0.0351	114	0.685	127	0.661	112
Turkish origin	0.0256	117	0.00877	114	0.126	127	0.152	112
Religious	0	117	0	114	0.449	127	0.420	112
Patriotic	0	117	0	114	0.165	127	0.214	112
Tradition-bound	0	117	0	114	0.0157	127	0.0268	112
Family-oriented	0.0342	117	0.0263	114	0.0866	127	0.125	112
Open	0.385	117	0.289	114	0.220	127	0.205	112
Conscientious	0.0598	117	0.0614	114	0.0709	127	0.0804	112
Extravert	0.393	117	0.421	114	0.441	127	0.384	112
Agreeable	0.427	117	0.404	114	0.283	127	0.357	112
Perception of pre-defined characteristics (5-point Likert scale)								
Christian	2.039	76	3.390	77	1.303	119	1.410	100
Muslim	3.934	76	2.125	72	4.672	122	4.545	101
German	3.434	99	3.873	102	2.982	109	3.449	98
Turkish	3.484	91	2.323	93	4.402	117	4.208	101
Italian	1.922	90	2.371	89	1.227	110	1.419	93
Polish	1.411	90	2.151	86	1.127	110	1.260	96
Punctual	3.278	72	3.381	84	3.385	65	2.933	60
Patriotic	2.107	75	2.282	85	4.196	107	4.184	98
Religious	2.778	72	2.643	84	4.413	121	4.260	104
Disciplined	3.430	86	3.495	91	3.382	89	3	81
Tradition-bound	2.974	77	2.975	81	4.209	115	4.048	104
Environmentally conscious	3.253	95	3.115	104	2.560	84	2.556	81
Others (5-point Likert scale)								
Callback	4.052	115	3.991	114	3.119	126	3.541	111
Attractivity	3.267	105	3.241	108	2.774	124	3.056	108
Realism	3.754	114	3.702	114	3.704	125	3.800	110

Note: The table reports means and number of observations for the second online pilot experiment depending on the treatment condition. The variables are dummy variables that equal 1 if respondents entered characteristics in an open-ended question that match the given category.

Table B.36: Second Online Pilot Experiment: Selected Summary Statistics

	Panel A: Ethnicity			
	Turkish	German	Ratio	Difference (p-value)
Turkish characteristics	0.385 (0.49)	0.345 (0.48)	1.12	-0.04 (0.368)
Perception: Turkish	4.000 (0.97)	3.304 (1.29)	1.21	-0.69*** (0.000)
Perception: Muslim	4.389 (0.81)	3.538 (1.45)	1.24	-0.85*** (0.000)
Callback	3.564 (1.12)	3.768 (1.08)	0.95	0.20** (0.039)
	Panel B: Minority Stereotypes			
	with Minority Stereotypes	without Minority Stereotypes	Ratio	Difference (p-value)
Turkish characteristics	0.674 (0.47)	0.048 (0.21)	14.04	-0.63*** (0.000)
Perception: Turkish	4.312 (0.81)	2.897 (1.11)	1.49	-1.42*** (0.000)
Perception: Muslim	4.614 (0.65)	3.054 (1.29)	1.51	-1.56*** (0.000)
Callback	3.316 (1.18)	4.022 (0.89)	0.82	0.71*** (0.000)
	Panel C: Stereotype conditions by ethnicity			
	Turkish			
	with Minority Stereotypes	without Minority Stereotypes	Ratio	Difference (p-value)
Turkish characteristics	0.685 (0.47)	0.059 (0.24)	11.61	-0.63*** (0.000)
Perception: Turkish	4.402 (0.78)	3.484 (0.96)	1.26	-0.92*** (0.000)
Perception: Muslim	4.672 (0.57)	3.934 (0.93)	1.19	-0.74*** (0.000)
Callback	3.119 (1.09)	4.052 (0.94)	0.77	0.93*** (0.000)
	German			
	with Minority Stereotypes	without Minority Stereotypes	Ratio	Difference (p-value)
Turkish characteristics	0.661 (0.48)	0.035 (0.18)	18.89	-0.63*** (0.000)
Perception: Turkish	4.208 (0.84)	2.323 (0.93)	1.81	-1.89*** (0.000)
Perception: Muslim	4.545 (0.74)	2.125 (0.92)	2.14	-2.42*** (0.000)
Callback	3.541 (1.24)	3.991 (0.85)	0.89	0.45** (0.013)
	Panel D: Ethnicity conditions by Stereotypes			
	with Minority Stereotypes			
	Turkish	German	Ratio	Difference (p-value)
Turkish characteristics	0.685 (0.47)	0.661 (0.48)	1.04	-0.02 (0.6896)
Perception: Turkish	4.402 (0.78)	4.208 (0.84)	1.05	-0.19* (0.0658)
Perception: Muslim	4.672 (0.57)	4.545 (0.74)	1.03	-0.13 (0.2607)
Callback	3.119 (1.09)	3.541 (1.24)	0.88	0.42*** (0.0032)
	without Minority Stereotypes			
	Turkish	German	Ratio	Difference (p-value)
Turkish characteristics	0.059 (0.24)	0.035 (0.18)	1.69	-0.02 (0.3784)
Perception: Turkish	3.484 (0.96)	2.323 (0.93)	1.49	-1.16*** (0.000)
Perception: Muslim	3.934 (0.93)	2.125 (0.92)	1.85	-1.81*** (0.000)
Callback	4.052 (0.94)	3.991 (0.85)	1.02	0.06 (0.3807)

Note: The table reports the means and standard deviations (in parentheses) of different treatment conditions. Turkish characteristics is a dummy variable that equals 1 if the respondent indicated Turkish characteristics in an open-ended question. Turkish, Muslim, and Callback are continuous variables measured on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Column 5 shows the differences in means and the respective p-values of a two-sample Wilcoxon rank-sum (Mann-Whitney) test of proportion testing that means are equal across conditions.

Table B.37: Second Online Pilot Experiment: Treatment Manipulation Checks

	(1) Perception: Turkish	(2) Perception: Turkish	(3) Perception: Muslim	(4) Perception: Muslim
Treatment: Turkish	1.258***	1.144***	1.978***	2.192***
	(0.141)	(0.411)	(0.144)	(0.374)
Treatment: Minority Stereotypes	1.944***	1.925***	2.621***	2.735***
	(0.140)	(0.416)	(0.131)	(0.350)
Treatment: Turkish × Minority Stereotypes	-1.107***	-1.084***	-1.923***	-1.890***
	(0.187)	(0.188)	(0.174)	(0.180)
Female	0.233**	0.236**	-0.0532	-0.0469
	(0.102)	(0.103)	(0.0808)	(0.0813)
Age	-0.00158	-0.00127	0.00485	0.00462
	(0.00478)	(0.00481)	(0.00453)	(0.00455)
Shared apartment	-0.156*	-0.154*	-0.0596	-0.0570
	(0.0925)	(0.0933)	(0.0814)	(0.0818)
German	0.293	0.248	0.226	0.463
	(0.313)	(0.458)	(0.294)	(0.379)
Turkish	-0.260	-1.186**	0.680***	0.323
	(0.552)	(0.549)	(0.245)	(0.297)
Treatment: Turkish × German		0.0856		-0.265
		(0.412)		(0.377)
Treatment: Turkish × Turkish		1.203		0.579
		(0.966)		(0.420)
Treatment: Minority Stereotypes × German		-0.00193		-0.139
		(0.419)		(0.342)
Treatment: Minority Stereotypes × Turkish		0.569		-0.355
		(0.898)		(0.423)
Foreign	0.373	0.372	0.0750	0.0697
	(0.251)	(0.253)	(0.244)	(0.248)
Instagram usage	0.0326	0.0387	0.0650**	0.0638**
	(0.0334)	(0.0345)	(0.0264)	(0.0272)
Constant	1.692***	1.675***	2.450***	2.212***
	(0.439)	(0.594)	(0.397)	(0.468)
Additional Survey Controls	Yes	Yes	Yes	Yes
Observations	362	362	336	336
R-squared	0.495	0.500	0.673	0.676

Note: The table reports multiple OLS models assessing subjects' perceptions of the individual's Turkish origin (columns 1–2) and Muslim affiliation (columns 3–4). Both dependent variables are measured on a five-point Likert scale ranging from "Do not agree at all" (1) to "completely agree" (5). Additional survey controls include a dummy variable indicating completion of the questionnaire, the time taken to complete it, and whether the participant had previously taken part in a similar study or held dual nationality. Robust standard errors in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.38: Second Online Pilot Experiment: Determinants of Callbacks

Callback	(1)	(2)	(3)
Treatment: Turkish	-0.118	0.0814	0.213
	(0.0964)	(0.120)	(0.382)
Treatment: Minority Stereotypes	-0.483***	-0.287*	0.0667
	(0.102)	(0.147)	(0.403)
Treatment: Turkish × Minority Stereotypes		-0.389**	-0.432**
		(0.198)	(0.197)
Attractivity	0.596***	0.586***	0.583***
	(0.0657)	(0.0670)	(0.0688)
Realism	0.0656	0.0535	0.0537
	(0.0643)	(0.0649)	(0.0656)
Female	0.0742	0.0774	0.0757
	(0.0986)	(0.0979)	(0.0984)
Age	-0.00776	-0.00832	-0.00796
	(0.00538)	(0.00531)	(0.00542)
Shared apartment	-0.0467	-0.0537	-0.0527
	(0.0969)	(0.0966)	(0.0968)
Turkish	-0.227	-0.245	0.101
	(0.507)	(0.519)	(0.345)
German	-0.625*	-0.683**	-0.449
	(0.321)	(0.312)	(0.373)
Treatment: Turkish × Turkish			-0.647
			(0.828)
Treatment: Turkish × German			-0.114
			(0.380)
Treatment: Minority Stereotypes × Turkish			1.468*
			(0.829)
Treatment: Minority Stereotypes × German			-0.367
			(0.406)
Foreign	-0.0511	-0.118	-0.128
	(0.257)	(0.248)	(0.250)
Instagram use	-0.0356	-0.0326	-0.0237
	(0.0313)	(0.0310)	(0.0317)
Constant	2.801***	2.945***	2.666***
	(0.496)	(0.501)	(0.535)
Additional Survey Controls	Yes	Yes	Yes
Observations	393	393	393
R-squared	0.305	0.312	0.319

Note: The table reports different OLS models using participants' stated intention to call back the fictitious profile owner as the dependent variable, measured on a five-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). Additional survey controls include a dummy variable for questionnaire completion, the time taken to complete the questionnaire, and whether the participant had previously participated in a similar study. Robust standard errors in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.7 Additional Online Survey

The purpose of the additional online survey was to gain subscribers/friends and likes on the images of the fictitious social media profiles in order to have profiles that would be difficult or impossible to detect as fictitious and to establish an initial social network. A small number of subscribers or friends could suggest a small social network or an inactive or fake account. The aforementioned is also applicable to a small number of likes on posts, which may suggest low popularity or relevance. This could also lead profile reviewers to believe that the account is inactive or fictitious.

However, the quality of subscribers/friends and likes on images and posts may also be relevant, not just the quantity. Participants who review the profiles in the experiments can see the list of subscribers/friends and theoretically review the applicant's friends (see Section 3.3.1.3). Although it is unlikely that someone receiving a friend request or searching for a new roommate and reviewing applications containing social media profiles would study the list of subscribers/friends of the applicant's profiles in great detail, it cannot be ruled out. Therefore, it is beneficial if subscribers/friends match the overall story of the applicant, i.e., having friends who are also students of a similar age at the same university.

Survey Design Firstly, we asked participants whether their Instagram profile was set to public or private. The list of subscriptions is publicly accessible only if the account is set to public. Secondly, we asked participants with public Instagram profiles to subscribe to only one of the experimental accounts in order to avoid detection of our fictitious experimental profiles. Participants with private Instagram profiles were randomly asked to subscribe to two of our four experimental accounts. Additionally, each participant was requested to subscribe randomly to a set of supplementary (private) experimental accounts. These accounts were used solely to act as close friends, for example, to tag them in photos or comment on posts. Participants were instructed to remain subscribed for a full year. Subsequently, we randomly requested participants to like some of the photos on the accounts (see Figure B.15).

Thirdly, we once again requested feedback on the perceived realism and authenticity of the profiles. Survey participants must have their own Instagram account to subscribe to the profile and like the posts. In contrast to the second online pilot, they are therefore able to access the entire profile without any restrictions. Therefore, we asked participants to rate the authenticity of the profiles and provide any remarks, suggestions, or comments in an open-ended question.

Results The additional online survey was conducted online between January and July 2023 in three waves using a professional survey platform. A total of $n = 528$ students were recruited. In contrast to the online pilot experiments, this survey was also available in English. 7.8% of participants completed the questionnaire in English. Approximately one-quarter of the participants indicated having a public profile.

In terms of perceived realism and authenticity of the profiles, the vast majority of respondents rated them extremely or very realistic (between 85.05% and 87.4%; see Table B.39). No respondent indicated that the profiles were unrealistic. The numbers are higher compared to the results of the second online pilot, presumably because participants can now view the entire profile's information and interactively access the profile on the social media platform, instead of examining screenshots.[9]

[9]Furthermore, we omitted certain details from the screenshots, such as the number of likes, which may have negatively affected the perceived realism.

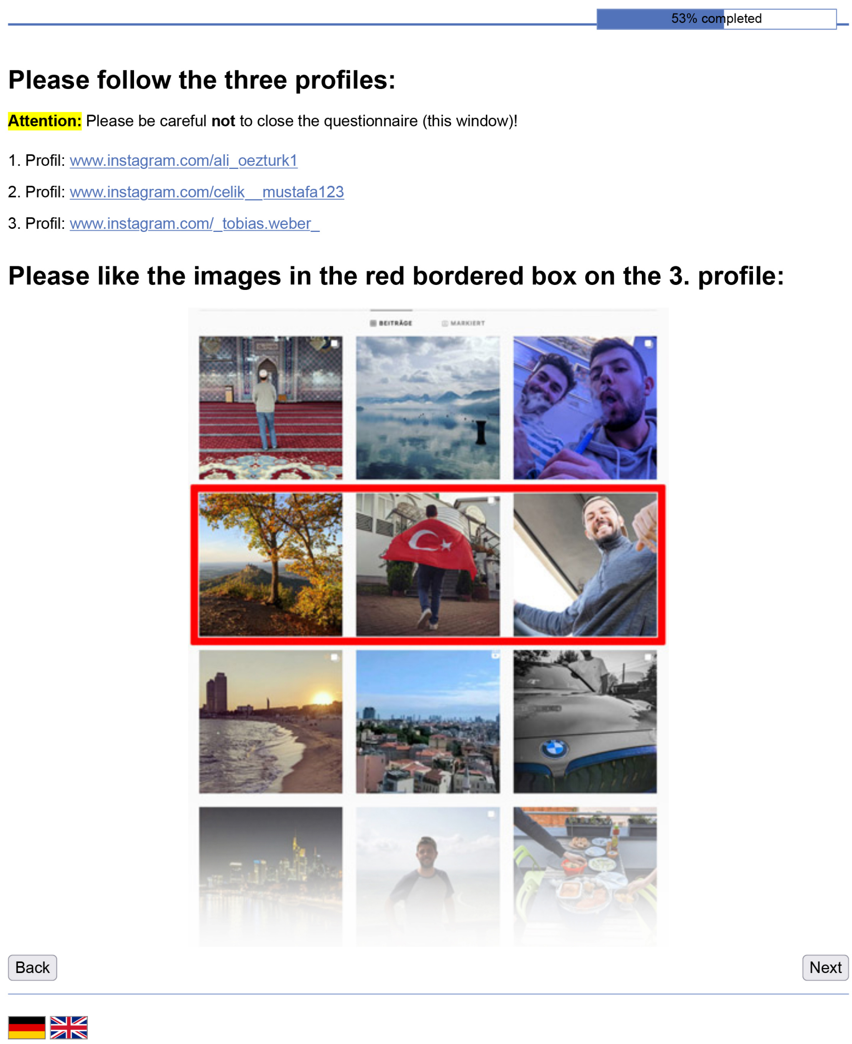


Figure B.15: Screenshot of the Follower Task (Additional Online Survey)

Table B.39: Additional Online Survey: Realism

	All	without Minority Stereotypes	with Minority Stereotypes	Turkish	German
Extremely realistic	36.56 [185]	28.36 [38]	39.52 [147]	36.22 [142]	37.64 [166]
Very realistic	50.20 [254]	56.72 [76]	47.85 [178]	50.26 [197]	49.66 [219]
Somewhat realistic	10.87 [55]	11.19 [15]	10.75 [40]	11.48 [45]	10.43 [46]
Only partly realistic	2.37 [12]	3.73 [5]	1.88 [7]	2.04 [8]	2.27 [10]
Not realistic at all	0 [0]	0 [0]	0 [0]	0 [0]	0 [0]

Note: The table reports the proportions of respondents who indicated the respective option for the subset of social media profiles provided. The numbers in brackets indicate the number of observations, i.e., the number of profiles that participants were presented with in the given condition.

B.8 Ethical Considerations

In field experiments on discrimination and correspondence tests especially, participants are not able to voluntarily decide whether they want to participate in the study. Informed consent cannot be obtained from participants due to the covert nature of the research. If participants know that they are part of an experiment, they often do not answer truthfully, particularly on sensitive topics such as discrimination (Zschirnt, 2019). Consequently, obtaining informed consent from participants would invalidate the results and hinder the exploration of the actual extent of discrimination in a particular market. Furthermore, our research inherently involves deceiving research subjects with fictitious applicants and social media users claiming to be existing individuals. Informing test subjects to avoid deception would again lead to a change in behavior, significantly reducing the validity of any results (Riach & Rich, 2004).

The following sections describe the potential costs for both participants and non-participants, as well as for the platforms involved. We then explain our efforts to minimize these costs and contrast the ethical considerations of our study to the social value of our research.

B.8.1 Costs to (Non-)Participants and Platforms

Direct Costs to Participants In the following, we distinguish between direct and indirect costs that participants may incur. Direct costs involve opportunity costs of time, including the time spent reviewing the fictitious friend request/application (including screening the social media information, depending on the treatment condition), making a decision, and accepting or declining the request or drafting and sending a response to the fictitious applicant.

Additionally, advertisers may require more time to find a suitable roommate. However, each ad in our data receives an average of 100 unique visits and about 14 applications within the first two hours of being online. Thus, the demand for rooms in shared apartments in Germany is comparatively high (MLP Finanzberatung SE, 2021), which limits the effect of our study on the average search duration in the housing market.

We aim to minimize the costs to the subjects of our study. Firstly, our experiment implements a non-matching pairs design, meaning that we send to each user/advertiser only one friend request/application. Consequently, each user/advertiser reviews a maximum of one request/application, instead of two or more, which is common in similar studies (Bertrand & Duflo, 2017; Evsyukova et al., 2025; Lippens et al., 2023; Zschirnt, 2019). In our study, discrimination cannot be detected at the individual level, as participants are faced with the decision to evaluate one potential friend/candidate only. Consequently, unequal treatment is only examined at the aggregate level.

Secondly, we implement a comprehensive duplicate check to ensure that each user/advertiser is only treated once. Thirdly, in Study II, we reject all responses within 24 to 48 hours, indicating that the fictitious applicant has already been offered another room on short notice (see Section 3.3.3). Rejecting a response earlier than 24 hours may suggest that the application was not authentic or that the response was not thoroughly read or acknowledged, potentially resulting in additional (indirect or psychological) costs for advertisers. Furthermore, some costs, particularly those related to communicating with a fictitious applicant, only apply to advertisers who respond to the application. The majority of advertisers tend to not response at all (see Section 3.4).

Indirect Costs to Participants In addition to the direct costs described above, participants may also incur indirect or psychological costs as a result of conducting the experiments. For instance, a rejection by the fictitious applicant after a positive response or invitation could cause disappointment or other emotions in the advertiser. This situation could be more serious if, for instance, the advertiser intentionally invited a person with an ethnic minority name or stereotypical images on their social media profile.

Furthermore, exposure to visual stereotypes has the potential to reinforce or perpetuate existing stereotypes, or alternatively, to create new or additional stereotypes. However, individuals may not interpret these stereotypes in the same way, and therefore they may not always have negative effects due to context dependency (Bordalo et al., 2016). For instance, one advertiser may reject religious practices and negatively connote religious stereotypes, while another may not. Additionally, we intentionally included the most frequently mentioned stereotypes in the images, so that the majority of advertisers are likely to associate these stereotypes with the given reference group, with or without exposure to visual stereotypes resulting from the experiment.

Another important concern is the privacy of advertisers, as they may not agree to some of their data being linked to their behavioral outcomes. However, we only access publicly available data without requiring registration or login. All the data we have collected is voluntarily published by the users/advertisers. Most users of the platform have chosen usernames that do not allow any conclusions to be drawn about their real names, which rules out the possibility of individual traceability. Additionally, we only analyze data on an aggregated level. We do not disclose any geographical data, including the addresses of shared apartments, even if they were publicly available in the ad. We did not use or store any information that was only available because a user accepted the friend request and is therefore not publicly available to everyone but to mutual connections only.

Costs to Non-Participants In addition to the costs incurred by participants of the experiment, there may also be costs for non-participants. For example, other friend requests or applications may be sorted out based on our fictitious application, either in terms of quality, if our request/application is perceived as a better fit, or if advertisers may no longer consider applications that are sent after a certain time or threshold regarding the number of received applications. Moreover, users or advertisers may unconsciously associate stereotypes presented by our fictitious applicants with real ethnic minority applicants, even if they do not conform to them, resulting in fewer invitations. Since we send our application or friend requests in large markets where individuals are exposed to significant amount of varying stereotypes, the overall negative impact on non-participants regarding stereotypical associations is expected to be limited, despite potential costs for some non-participants, which may not be negligible.

Costs to the Platforms The fictitious social media profiles were created on the photo- and video-sharing platform "Instagram" which was acquired by Facebook (today Meta Platforms) in 2012. As of 2021, Instagram had approximately 1.21 billion users worldwide (Insider Intelligence, 2022). Based on the high number of Instagram users and profiles, we argue that the costs in terms of server capacity and/or user base are negligible. Additionally, we have created only a small number of fictitious profiles in comparison to the large number of Instagram users, ensuring that the platform's perceived quality remains unaffected. Although an increased number of fictitious profiles may affect the perceived quality of a social media platform, we posit that our profiles are not perceived as such. This is supported by the fact that they were predominantly rated as highly realistic by participants in the online pilot studies.

The fictitious Instagram profiles were used to apply for vacant rooms on the largest website for shared apartment ads in Germany. The website had approximately 17 million visitors in August 2022.[10] We created six fictitious accounts on this platform. Although we do not have data on account numbers, we believe that the costs of server capacity and user base are negligible, given the high number of visits on the platform. This ensures that the perceived quality of the platform remains unaffected.

In addition, as soon as we encountered obvious bot or spam profiles or spam ads, we immediately reported them to the respective platform, which also increases the perceived platform quality and provides an important service.

Overall, we believe that the costs for platforms are negligible. Given that online platforms have a significant impact on real-world outcomes, conducting scientific studies on social media and other platforms is essential to understand the wide-ranging impact of platforms in general.

Absence of Debriefing Debriefing, i.e., informing the research subjects ex-post about participating in an experiment and the purpose of it, is a relevant tool in order to mitigate the negative consequences of deception and voluntary consent. However, debriefing participants

[10] See: https://www.wg-gesucht.de/ic/online-werbung.html [Retrieved: February 8, 2024].

may result in high costs for the research subjects. Firstly, there are time costs associated with reading the debriefing. Secondly, participants may experience significant psychological costs if they feel that they have behaved in a potentially discriminatory manner. Thirdly, there may also be costs for non-participants, as debriefed advertisers may anticipate a fictitious person in future applications from individuals with an ethnic minority name, resulting in a lower response rate to such applications. Therefore, we decided against debriefing participants. In our opinion, the disadvantages, which include significant additional costs for both participants and non-participants, outweigh the advantages.

B.8.2 Benefits

Our study contributes to the understanding of discrimination and the impact of stereotypes and social media information on unequal treatment. To date, there are no causal studies on the effect of visual stereotypes on unequal treatment. We fill this research gap by providing causal evidence on the role of social media information, specifically visual stereotypes, in affecting informal social network formation and market access for minority groups. In addition to the ethical implications of field experiments and correspondence tests in general and our identification strategy in particular, our study provides important implications for individuals affected by discrimination, as we show how their appearance on social media significantly influence unequal treatment.

Our results provide important evidence for policymakers to better understand the role of stereotypical social media information in the intent to discriminate, which can improve and enhance anti-discrimination laws which rather focus on professional than non-professional, informal settings. Furthermore, the results of our study may foster the public discourse on the impact of online information on disadvantaged groups, marginalized communities, and ethnic minority members in general.

Additionally, social media has a significant impact on states, economies, and individuals by enabling political participation and resistance, as well as influencing social connections, well-being, mental health, and information acquisition of individual users (Acquisti et al., 2015; Allcott et al., 2020; Allcott & Gentzkow, 2017; Ekman, 2019). Therefore, experimental studies on the impacts of social media are vital to enhance our understanding of its effects on our daily lives and how online behavior affect offline outcomes.

B.8.3 Assessment of Costs and Benefits in the Employed Setting

As discrimination research is crucial for society (Bursell, 2007; Pager, 2007; Zschirnt, 2019), it is necessary to investigate the extent of discrimination, which factors determine discriminatory intentions of market participants, and the behavior or information that promote *or* impede unequal treatment.

Our study is the first to present causal evidence on how different social media contents of the same user/applicant affect informal social network formation and selection decisions. We examine how visual information that conforms to *or* contradicts ethnic minority stereotypes affects unequal treatment. Furthermore, due to the prevalence of misinformation on social media, particularly targeting ethnic minorities and immigrants, it is important to understand the impact of ethnic stereotypes and the selection of ethnic minority applicants.

The selected setting is necessary and the least invasive to study the effects of visual stereotypes on discrimination in personal and social settings. Although observational data would reduce the costs to participants, this alternative does not allow for a comprehensive investigation of discrimination in our context. While representative samples are desirable, they do not facilitate causal studies, which limits our ability to examine the direct effect of social media information on discrimination. Laboratory studies are susceptible to biases such as the Hawthorne effect and experimenter demand bias, which can make it less likely to yield externally valid results.

Moreover, visual stereotypes may include characteristics that some individuals perceive as negative, which can complicate the interpretation of findings when participants are aware that they are part of an experiment and may provide socially desirable answers. However, due to the potential impact of social media information and visual stereotypes on individuals and market access, a field experiment or correspondence test is necessary in our context, despite the costs incurred by participants.

Legislation has recognized the need for deception in field experiments, as the disclosure of information would impair the validity of the results (Zschirnt, 2019). In addition, field experiments on discrimination provide necessary evidence for the enforcement of anti-discrimination laws.

Although our research design imposes costs on participants and has important ethical implications to consider, the societal value of our research and the importance of the research question justify these costs. Additionally, as described above, we have taken several steps to minimize costs for both participants, and non-participants, and platforms. In conclusion, we assert that the advantages of obtaining causal evidence of discrimination and the moderating role of various social media information through a field experiment outweigh the costs imposed on (non-)participants.

Appendix C

Appendix Chapter 4

C.1 Tables & Figures

Table C.1: Summary Statistics

		Mean	S.D.	Obs.
Experimental Variables				
Ethnic minority	(1/0)	0.488	0.500	4,212
Instagram account	(1/0)	0.667	0.471	4,212
Premium account	(1/0)	0.495	0.500	4,212
Minority stereotypes	(1/0)	0.333	0.471	4,212
Treatment: SM-Profile with minority stereotypes	(1/0)	0.333	0.471	4,212
Treatment: SM-Profile without minority stereotypes	(1/0)	0.334	0.472	4,212
Treatment: Without SM-Profile	(1/0)	0.333	0.471	4,212
Number of applications per name and week	(#)	234.134	71.585	4,212
Number of SM applications per name and week	(#)	130.712	55.986	4,212
1st week (wave 1)	(1/0)	0.120	0.325	4,212
2nd week (wave 1)	(1/0)	0.117	0.321	4,212
3rd week (wave 1)	(1/0)	0.142	0.349	4,212
4th week (wave 1)	(1/0)	0.117	0.321	4,212
1st week (wave 2)	(1/0)	0.151	0.358	4,212
2nd week (wave 2)	(1/0)	0.031	0.173	4,212
3rd week (wave 2)	(1/0)	0.067	0.250	4,212
4th week (wave 2)	(1/0)	0.063	0.244	4,212
5th week (wave 2)	(1/0)	0.015	0.122	4,212
6th week (wave 2)	(1/0)	0.000	0.000	4,212
7th week (wave 2)	(1/0)	0.000	0.000	4,212
8th week (wave 2)	(1/0)	0.081	0.272	4,212
9th week (wave 2)	(1/0)	0.097	0.295	4,212
Aachen	(1/0)	0.048	0.214	4,212
Berlin	(1/0)	0.180	0.385	4,212
Bochum	(1/0)	0.024	0.154	4,212
Darmstadt	(1/0)	0.036	0.187	4,212
Dresden	(1/0)	0.042	0.202	4,212
Düsseldorf	(1/0)	0.033	0.178	4,212
Frankfurt am Main	(1/0)	0.067	0.250	4,212
Gießen	(1/0)	0.029	0.168	4,212
Göttingen	(1/0)	0.036	0.186	4,212
Hamburg	(1/0)	0.094	0.292	4,212
Köln	(1/0)	0.066	0.249	4,212

Continued on next page

Table C.1 – continued from previous page

		Mean	S.D.	Obs.
Leipzig	(1/0)	0.053	0.224	4,212
München	(1/0)	0.121	0.326	4,212
Münster	(1/0)	0.045	0.208	4,212
Stuttgart	(1/0)	0.124	0.330	4,212
Response Variables				
Response	(1/0)	0.441	0.497	4,212
Callback	(1/0)	0.302	0.459	4,212
Rejection	(1/0)	0.059	0.236	4,212
Other response	(1/0)	0.079	0.270	4,212
Callback or other Response	(1/0)	0.381	0.486	4,212
Time between application and response	(days)	2.524	5.039	1,856
2nd message after initial message	(1/0)	0.218	0.413	1,856
Response received outside of platform	(1/0)	0.006	0.077	1,856
Number of characters in response text	(#)	245.683	219.159	1,856
Number of smileys and emojis in response text	(#)	0.252	0.557	1,856
Room & Shared Apartment Characteristics				
Roomsize	(sqm)	17.122	7.835	4,211
Total monthly rent	(€)	550.253	186.070	4,212
Monthly rent	(€)	476.726	183.957	4,212
Additional monthly costs (utilities)	(€)	73.832	61.036	3,777
Other monthly costs	(€)	10.698	25.171	2,882
Deposit	(€)	844.815	564.991	3,869
Clearance Payment	(€)	83.055	269.129	2,761
Pictures included	(1/0)	0.966	0.181	4,212
Temporary	(1/0)	0.143	0.350	4,212
Availability (if temporary)	(days)	389.125	293.527	602
Online time	(minutes)	119.289	90.573	4,212
Number of vacant rooms	(#)	1.267	0.930	4,212
Apartment size	(sqm)	93.094	48.869	3,139
Smoking permitted	(1/0)	0.482	0.500	3,069
Online viewing possible	(1/0)	0.364	0.481	4,212
Ad text states that roommates do not discriminate	(1/0)	0.007	0.083	4,212
Applicant should mention codeword(s)	(1/0)	0.035	0.184	4,212
Applicant is asked to include social media profile	(1/0)	0.098	0.298	4,212
Number of smileys and emojis in ad text	(#)	1.473	2.677	4,212
Length of ad text	(#)	2062.214	1387.108	4,212
Only males accepted	(1/0)	0.061	0.239	4,212
All gender accepted	(1/0)	0.939	0.239	4,212
Age preference for applicants	(1/0)	0.622	0.485	4,212
Roommate Characteristics				
Female roommates	(#)	0.819	0.976	4,212
Male roommates	(#)	1.205	1.134	4,212
Diverse roommates	(#)	0.020	0.171	4,212
Mixed gender	(1/0)	0.491	0.500	4,212
Average age	(years)	26.675	6.113	3,156
Number of languages spoken by the roommates	(#)	2.227	1.131	3,522
Roommates speak German	(1/0)	0.816	0.387	4,212
Roommates speak English	(1/0)	0.671	0.470	4,212
Roommates speak Turkish	(1/0)	0.023	0.150	4,212
Roommates speak Arabic	(1/0)	0.014	0.119	4,212
Roommates speak more than two languages	(1/0)	0.275	0.446	3,522
Students	(1/0)	0.613	0.487	4,212

Continued on next page

Table C.1 – continued from previous page

		Mean	S.D.	Obs.
Communally	(1/0)	0.428	0.495	4,212
Non-communally	(1/0)	0.097	0.296	4,212
Males only	(1/0)	0.041	0.199	4,212
Females only	(1/0)	0.059	0.235	4,212
Young professionals	(1/0)	0.062	0.242	4,212
Employed	(1/0)	0.453	0.498	4,212
Students' hall of residence	(1/0)	0.013	0.113	4,212
Vegetarians/vegans	(1/0)	0.050	0.217	4,212
Cross-generational)	(1/0)	0.026	0.158	4,212
Single mother/father	(1/0)	0.007	0.085	4,212
With children	(1/0)	0.013	0.111	4,212
Fraternity	(1/0)	0.062	0.242	4,212
LGBTQIA roommates	(1/0)	0.089	0.285	4,212
Elderly roommates	(1/0)	0.005	0.070	4,212
Disabled roommates	(1/0)	0.026	0.159	4,212
Shared apartment is/will be newly established	(1/0)	0.065	0.246	4,212
Internationals welcome	(1/0)	0.148	0.355	4,212
Advertiser Characteristics				
Female	(1/0)	0.403	0.491	4,087
Age of advertiser	(years)	32.771	12.645	1,841
German origin	(1/0)	0.665	0.472	4,212
Turkish origin	(1/0)	0.018	0.131	4,212
Arabic origin	(1/0)	0.029	0.169	4,212
Origin from muslim-majority country	(1/0)	0.079	0.270	4,212
Account age	(months)	50.918	42.624	4,211
Profile picture	(1/0)	0.509	0.500	4,212
Premium account	(1/0)	0.006	0.077	1,854
Premium Variables & Ad Statistics				
Favorites	(#)	5.882	9.294	4,212
Visits	(#)	116.482	141.308	4,212
Applications	(#)	16.212	25.050	4,212
Applications from premium accounts	(#)	1.772	3.934	4,212
Share of applications from premium accounts	(%)	8.215	13.796	4,212
Position in advertiser's inbox (non-premium appl.)	(#)	16.353	24.186	2,129
Position in advertiser's inbox (premium appl.)	(#)	2.886	4.081	2,083
Responses from advertiser	(#)	2.575	6.806	2,125
Average time of advertiser's response	(hours)	16.366	23.047	1,080
Favorites after 24h	(#)	10.507	13.130	1,943
Visits after 24h	(#)	205.464	186.988	1,943
Applications after 24h	(#)	29.314	36.004	1,943
Applications from premium accounts after 24h	(#)	4.373	6.565	1,943
Share of applications from premium accounts after 24h	(%)	17.346	18.186	1,943
Responses from advertiser after 24h	(#)	4.051	6.822	1,943
Average time of advertiser's response after 24h	(hours)	13.093	15.820	1,365
Ad deactivated after 24h	(1/0)	0.064	0.244	2,075
Geographic Variables				
Out of town	(1/0)	0.019	0.135	4,212
Distance to city center	(kilometers)	4.619	3.500	4,209
Distance to district center	(kilometers)	2.686	2.196	3,713
Distance to main train station	(kilometers)	4.656	3.460	4,209
Bars in 500m radius	(#)	8.314	7.423	4,212
Churches in 1km radius	(#)	10.972	6.320	4,212

Continued on next page

Table C.1 – continued from previous page

		Mean	S.D.	Obs.
Mosques in 1km radius	(#)	1.298	2.058	4,212
Mosques in 3km radius	(#)	6.844	5.363	4,212
University/faculty buildings in 1km radius	(#)	7.874	7.667	4,212
Distance to largest university in city	(kilometers)	6.046	5.098	4,211
Distance to 2nd largest university in city	(kilometers)	5.258	3.788	4,211
Distance to 3rd largest university in city	(kilometers)	6.695	5.500	3,537
Avg. distance to largest two universities in city	(kilometers)	5.652	4.111	4,211
District Population Variables				
Population	(#)	144254.765	123837.291	4,134
Female population	(#)	72452.506	62473.629	4,134
Share of female population	(%)	0.501	0.017	4,134
Population density	(#)	5043.268	3171.315	4,015
Foreign population	(#)	35425.152	33882.251	4,134
Share of foreign population	(#)	0.248	0.083	4,134
Turkish population	(#)	5616.431	8851.899	4,132
Share of Turkish population	(%)	0.030	0.024	4,132
Share of Turkish / Foreign population	(%)	0.112	0.078	4,132
Households	(#)	58706.260	50168.958	3,257
Apartments	(#)	90108.176	70037.347	3,071
Unemployment rate	(%)	6.861	2.637	3,529
Unemployed	(#)	6498.787	6726.901	4,015
Average age	(years)	41.551	2.049	4,015
Share of migrants	(%)	0.265	0.116	3,744
Rent control (“Mietpreisbremse”) in city	(1/0)	0.898	0.303	4,134
Population change 2017-2021 (AGR)	(%)	0.004	0.008	4,134
Foreign population change 2017-2021 (AGR)	(%)	0.036	0.029	4,134
Turkish population change 2017-2021 (AGR)	(%)	0.021	0.052	4,130

Note: The table presents summary statistics on experimental, response, geographic, district population, and premium variables, as well as room & shared apartment, roommate, and advertiser characteristics.

Table C.2: Summary Statistics of Experimental, Response, and Ad Statistic Variables by Experimental Waves

		First Wave Mean/SD	**Second Wave** Mean/SD	Difference
Experimental Variables				
Ethnic minority	(1/0)	0.468 (0.499)	0.508 (0.500)	-0.040**
Instagram account	(1/0)	0.667 (0.472)	0.668 (0.471)	-0.001
Premium account	(1/0)	0.488 (0.500)	0.501 (0.500)	-0.013
Minority stereotypes	(1/0)	0.334 (0.472)	0.331 (0.471)	0.003
Treatment: Without SM-Profile	(1/0)	0.333 (0.472)	0.332 (0.471)	0.001
Treatment: SM-Profile without minority stereotypes	(1/0)	0.332 (0.471)	0.336 (0.473)	-0.004
Treatment: SM-Profile with minority stereotypes	(1/0)	0.334 (0.472)	0.331 (0.471)	0.003
Response Variables				
Response	(1/0)	0.451 (0.498)	0.430 (0.495)	0.021
Callback	(1/0)	0.314 (0.464)	0.290 (0.454)	0.024
Rejection	(1/0)	0.057 (0.231)	0.062 (0.241)	-0.006
Other response	(1/0)	0.080 (0.272)	0.078 (0.268)	0.002
Callback or other Response	(1/0)	0.395 (0.489)	0.368 (0.482)	0.027
Time between application and response	(days)	2.722 (5.490)	2.319 (4.521)	0.403
2nd message after initial message	(1/0)	0.053 (0.224)	0.049 (0.216)	0.004
Premium Variables & Ad Statistics				
Favorites	(#)	4.979 (8.338)	6.769 (10.070)	-1.790***
Visits	(#)	96.511 (106.601)	136.096 (166.286)	-39.585***
Applications	(#)	13.362 (20.798)	19.010 (28.346)	-5.648***
Applications from premium accounts	(#)	0.738 (1.586)	2.786 (5.112)	-2.048***
Share of applications from premium accounts	(%)	4.365 (9.557)	11.995 (16.086)	-7.630***
Position in advertiser’s inbox (non-premium)	(#)	13.852 (21.034)	18.875 (26.768)	-5.023***
Position in advertiser’s inbox (premium)	(#)	1.759 (1.566)	3.963 (5.279)	-2.204***
Obs.		2,087	2,125	

Note: The table presents sub-sample summary statistics for the experimental, response, and ad statistic variables across the first and second experimental waves. The first wave was conducted between October and November 2023, while the second wave took place between September and November 2024. Column 5 reports the results of the two-sample t-test. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.3: Summary Statistics of Social Media Variables

	Ethnic Minority Without Minority Stereotypes Mean/SD	**Ethnic Minority With Minority Stereotypes** Mean/SD
Weekly Followers	211.104 (18.973)	220.387 (1.562)
Weekly Following/Subscriptions	329.025 (53.410)	323.253 (57.553)
Weekly Visits	21.730 (10.005)	27.836 (14.914)
Weekly Impressions	198.572 (72.303)	201.079 (151.422)
Weekly Reach	14.110 (6.418)	55.933 (83.981)
Weekly Engagement	0.000 (0.000)	0.927 (2.462)
Weekly Visits per Application	0.150 (0.063)	0.190 (0.081)
Weekly Impressions per Application	1.365 (0.524)	1.281 (0.809)
Weekly Reach per Application	0.098 (0.044)	0.403 (0.611)
Weekly Engagement per Application	0.000 (0.000)	0.007 (0.017)
Obs.	670	683
	Ethnic Majority Without Minority Stereotypes Mean/SD	**Ethnic Majority With Minority Stereotypes** Mean/SD
Weekly Followers	216.343 (12.479)	207.238 (5.261)
Weekly Following/Subscriptions	321.671 (76.844)	313.693 (58.722)
Weekly Visits	35.752 (9.251)	40.014 (13.326)
Weekly Impressions	290.278 (97.644)	299.955 (118.948)
Weekly Reach	20.660 (6.020)	24.332 (8.631)
Weekly Engagement	0.598 (1.607)	0.551 (1.502)
Weekly Visits per Application	0.232 (0.074)	0.270 (0.141)
Weekly Impressions per Application	1.859 (0.671)	1.912 (0.711)
Weekly Reach per Application	0.135 (0.051)	0.159 (0.069)
Weekly Engagement per Application	0.004 (0.010)	0.007 (0.019)
Obs.	727	730

Note: The table reports summary statistics of social media variables on the number of followers, subscriptions, visits, impressions, reach, and engagement for the respective profiles/conditions. Profile visits are defined as the number of times a profile was accessed (i.e., unique visits) by a user who viewed a post or story (see: https://help.instagram.com/825941707897287 [Retrieved: January 20, 2025]). Impressions are defined as the number of times a post or story was on screen (see: https://help.instagram.com/825941707897287 [Retrieved: January 20, 2025]). Reach is defined as the number of unique users that have seen posts or stories at least once. It is different from impressions, which may include multiple views by the same user (see: https://help.instagram.com/788388387972460 and https://help.instagram.com/825941707897287 [Retrieved: January 20, 2025]). Engagement is defined as the number of interactions users take when they engage with the contents, such as likes, comments, saves and shares (see: https://help.instagram.com/788388387972460 [Retrieved: January 20, 2025]).

Figure C.1: Mean Callback Rates for High and Low Information Salience and Information Conditions

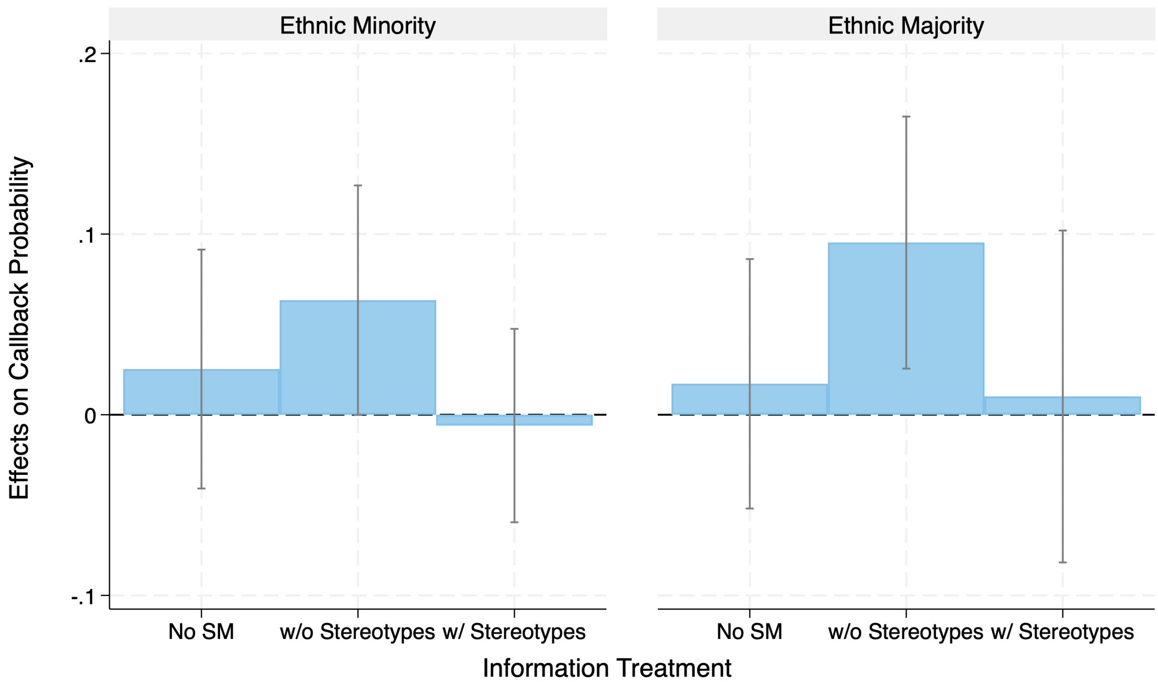


Figure C.2: Average Marginal Effects of Higher Information Salience

Table C.4: Mean Callback Rates

	Low Salience	**High Salience**	**Diff.**	**Ratio**
All Applications	28.69	31.73	3.05**	1.11
	[2,129]	[2,083]	(0.032)	
Panel A: Both Ethnicities				
Ethnic Minority	21.29	24.14	2.85	1.13
	[1,038]	[1,019]	(0.123)	
Ethnic Majority	35.75	39.00	3.25	1.09
	[1,091]	[1,064]	(0.118)	
Panel B: Ethnic Minority				
No SM-Profile	21.70	25.00	3.30	1.15
	[364]	[340]	(0.302)	
SM-Profile Without Minority Stereotypes	22.35	28.24	5.89*	1.26
	[349]	[340]	(0.076)	
SM-Profile With Minority Stereotypes	19.69	19.17	-0.52	0.97
	[325]	[339]	(0.866)	
Panel C: Ethnic Majority				
No SM-Profile	35.81	37.01	1.20	1.03
	[363]	[335]	(0.742)	
SM-Profile Without Minority Stereotypes	39.94	46.99	7.05*	1.18
	[353]	[366]	(0.057)	
SM-Profile With Minority Stereotypes	31.73	32.78	1.05	1.03
	[375]	[363]	(0.761)	

Note: The table reports mean callback rates for the different experimental conditions. The numbers in brackets in each cell present the number of applications sent for the given sub-sample. Column 5 shows the p-value for a two-sample Wilcoxon rank-sum (Mann-Whitney) test of proportion testing the null hypothesis that the callback rates are equal across low and high information salience. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.5: The Effects of Higher Salience & Information Treatments on Callback Probability

Callback	(1)	(2)
High Salience	0.0203	0.0206
	(0.0256)	(0.0247)
Ethnic Minority	-0.139***	-0.142***
	(0.0200)	(0.0184)
Treatment 1: Without SM (Ref.)	-	-
Treatment 2: SM Without Minority Stereotypes	0.0190	0.0212
	(0.0193)	(0.0136)
Treatment 3: SM With Minority Stereotypes	-0.0357	-0.0437*
	(0.0227)	(0.0227)
High Salience × Treatment 1 (Ref.)	-	-
High Salience × Treatment 2	0.0701**	0.0712**
	(0.0348)	(0.0328)
High Salience × Treatment 3	-0.0117	-0.0110
	(0.0470)	(0.0395)
Ethnic Minority × High Salience × Treatment 1 (Ref.)	-	-
Ethnic Minority × High Salience × Treatment 2	-0.0332	-0.0483
	(0.0400)	(0.0372)
Ethnic Minority × High Salience × Treatment 3	-0.0116	0.00179
	(0.0264)	(0.0306)
Observations	4,212	4,130
Pseudo R^2	0.0610	0.147
Week & Wave FE	Yes	Yes
City FE	Yes	Yes
All Additional Controls	No	Yes
P-value Heteroscedasticity (Wald) Test	0.183	0.228

Note: The table reports the average marginal effects computed from different probit models with callback as the dependent variable. All additional controls include room & shared apartment, roommate, advertiser, ad statistic, geographic, and district level demographic controls (see notes to Table 4.1 for a detailed description. Robust standard errors (in parentheses) are clustered at the city level and robust to heteroscedasticity. The Wald test yields $\chi^2 = 1.77$ ($p = 0.183$) for model 1 and $\chi^2 = 1.45$ ($p = 0.228$) for model 2, indicating no significant heteroscedasticity (Heckman, 1998; Neumark, 2012). * p < 0.1, ** p < 0.05, *** p < 0.01.

Table C.6: The Effects of Higher Salience on Callback Probability (Split Samples for Low vs. High Salience)

Callback	(1) Low Salience	(2) Low Salience	(3) High Salience	(4) High Salience
Inbox Position	-0.00740***	-0.00572***	-0.0422***	-0.0358***
	(0.00112)	(0.00180)	(0.00570)	(0.00554)
Ethnic Minority	-0.144***	-0.128***	-0.130***	-0.136***
	(0.0284)	(0.0306)	(0.0421)	(0.0373)
Treatment 1: Without SM (Ref.)	-	-	-	-
Treatment 2: SM Without Minority Stereotypes	0.0406	0.0380	0.0976***	0.0981***
	(0.0418)	(0.0299)	(0.0252)	(0.0280)
Treatment 3: SM With Minority Stereotypes	-0.0658	-0.0504*	-0.0408	-0.0498
	(0.0420)	(0.0298)	(0.0360)	(0.0308)
Ethnic Minority × Treatment 1 (Ref.)	-	-	-	-
Ethnic Minority × Treatment 2	-0.0350	-0.0451	-0.0532	-0.0580
	(0.0510)	(0.0436)	(0.0500)	(0.0502)
Ethnic Minority × Treatment 3	0.0387	0.00891	-0.0174	-0.00633
	(0.0384)	(0.0294)	(0.0448)	(0.0456)
Observations	2,129	2,098	2,083	2,032
Pseudo R^2	-	-	0.100	0.172
Week & Wave FE	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
All Additional Controls	No	Yes	No	Yes
P-value Heteroscedasticity (Wald) Test	0.000	0.053	0.372	0.992

Note: The table reports the average marginal effects computed from different (heteroskedastic) probit models with callback as the dependent variable. All additional controls include room & shared apartment, roommate, advertiser, ad statistic, geographic, and district level demographic controls (see notes to Table 4.1 for a detailed description. Robust standard errors (in parentheses) are clustered at the city level. Column 1 and 2 estimate average marginal effects of heteroskedastic probit models, while columns 3 and 4 estimate average marginal effects of regular probit models. The Wald test yields $\chi^2 = 15.11$ ($p = 0.000$) for model 1, $\chi^2 = 3.75$ ($p = 0.053$) for model 2, $\chi^2 = 0.80$ ($p = 0.372$) for model 3, and $\chi^2 = 0.00$ ($p = 0.992$) for model 4, indicating no significant heteroscedasticity for models 3 and 4 (Heckman, 1998; Neumark, 2012). * p < 0.1, ** p < 0.05, *** p < 0.01.

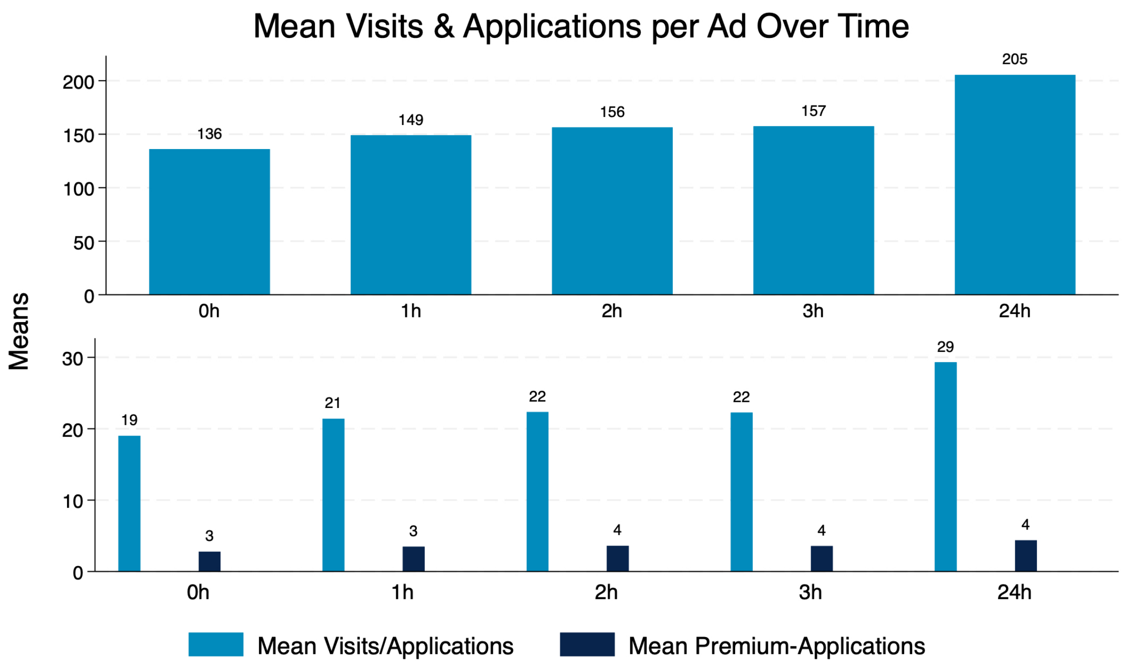


Figure C.3: Mean Visits & Applications per Ad Over Time (After Sending out the Application; Second Wave)

Table C.7: The Effects of Inbox Position on Callback Probability

Callback	(1)	(2)	(3)	(4)
Inbox Position	-0.00272***	-0.00271***	-0.00647***	-0.00645***
	(0.00101)	(0.00100)	(0.000911)	(0.000915)
Inbox Position^2	-	-	3.43e-05***	3.42e-05***
			(3.85e-06)	(3.92e-06)
Ethnic Minority	-0.130***	-0.112***	-0.137***	-0.120***
	(0.0143)	(0.0252)	(0.0134)	(0.0258)
Ethnic Minority × Inbox Position	-0.00374***	-0.00372***	-0.00246**	-0.00244**
	(0.00135)	(0.00133)	(0.00100)	(0.000987)
High Salience	-0.00559	-0.00551	-0.0239	-0.0239
	(0.0153)	(0.0154)	(0.0157)	(0.0157)
Treatment 1: Without SM (Ref.)	-	-	-	-
Treatment 2: SM Without Minority Stereotypes	0.0456***	0.0650***	0.0450***	0.0641***
	(0.0132)	(0.0116)	(0.0133)	(0.0117)
Treatment 3: SM With Minority Stereotypes	-0.0508***	-0.0468**	-0.0503***	-0.0461**
	(0.0140)	(0.0204)	(0.0140)	(0.0206)
Ethnic Minority × Treatment 1 (Ref.)	-	-	-	-
Ethnic Minority × Treatment 2	-	-0.0434	-	-0.0425
		(0.0359)		(0.0357)
Ethnic Minority × Treatment 3	-	-0.00770	-	-0.00816
		(0.0286)		(0.0290)
Observations	4,130	4,130	4,130	4,130
Pseudo R^2	0.151	0.152	0.154	0.155
Week & Wave FE	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
All Additional Controls	Yes	Yes	Yes	Yes
P-value Heteroscedasticity (Wald) Test	0.918	0.803	0.928	0.699

Note: The table reports similar specifications as in Table 4.1 including the position of the application in the advertiser's inbox. A smaller number indicates a higher inbox position. All additional controls include room & shared apartment, roommate, advertiser, ad statistic, geographic, and district level demographic controls (see notes to Table 4.1 for a detailed description. Robust standard errors (in parentheses) are clustered at the city level and robust to heteroscedasticity. The Wald test yields $\chi^2 = 0.01$ ($p = 0.918$) for model 1, $\chi^2 = 0.06$ ($p = 0.803$) for model 2, $\chi^2 = 0.01$ ($p = 0.928$) for model 3, and $\chi^2 = 0.15$ ($p = 0.699$) for model 4, indicating no significant heteroscedasticity (Heckman, 1998; Neumark, 2012). * p < 0.1, ** p < 0.05, *** p < 0.01.

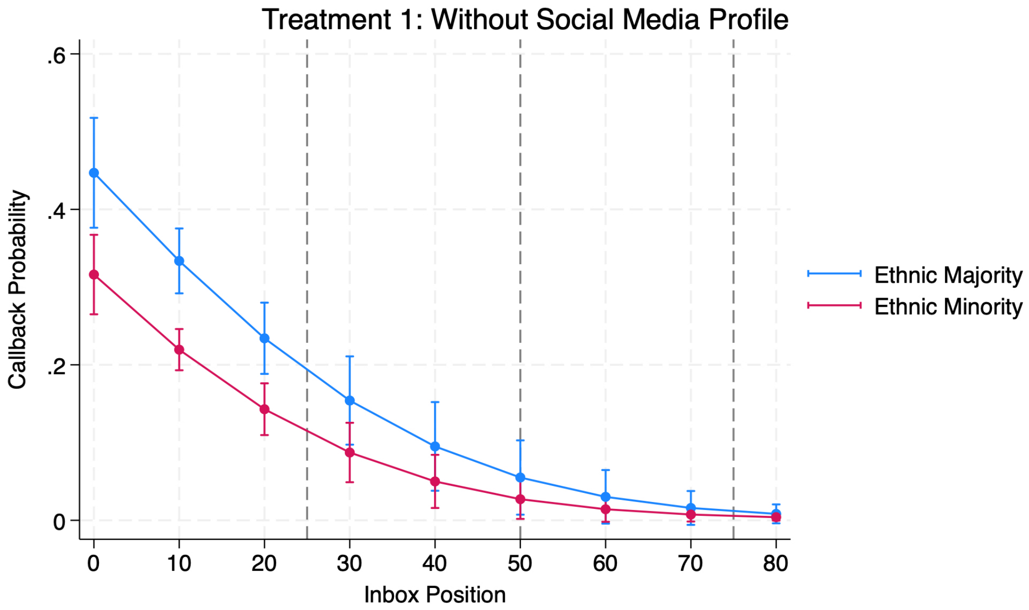


Figure C.4: Predicted Callback Probability by Inbox Position and Ethnicity (Treatment 1)

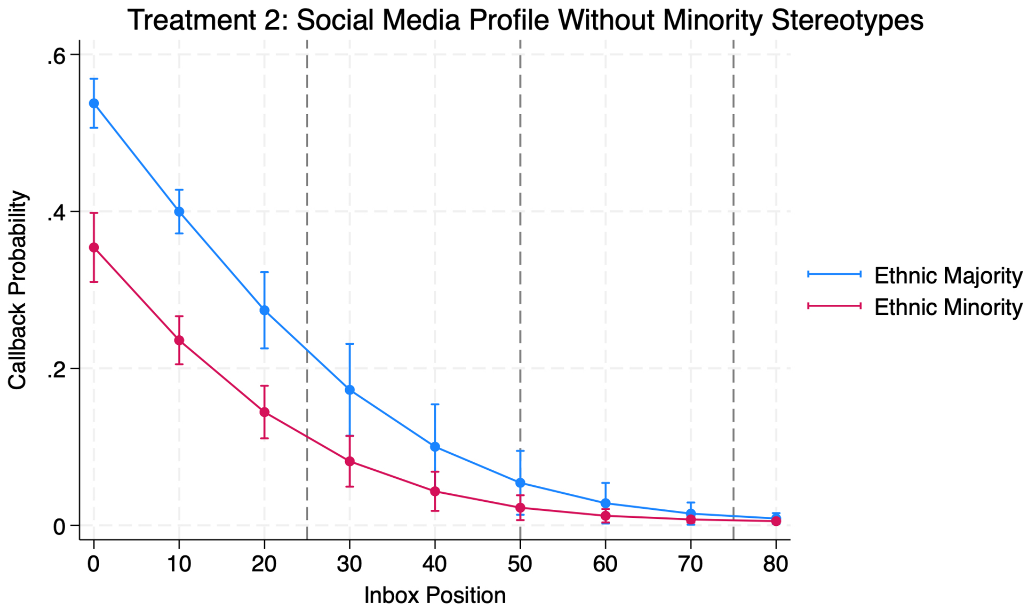


Figure C.5: Predicted Callback Probability by Inbox Position and Ethnicity (Treatment 2)

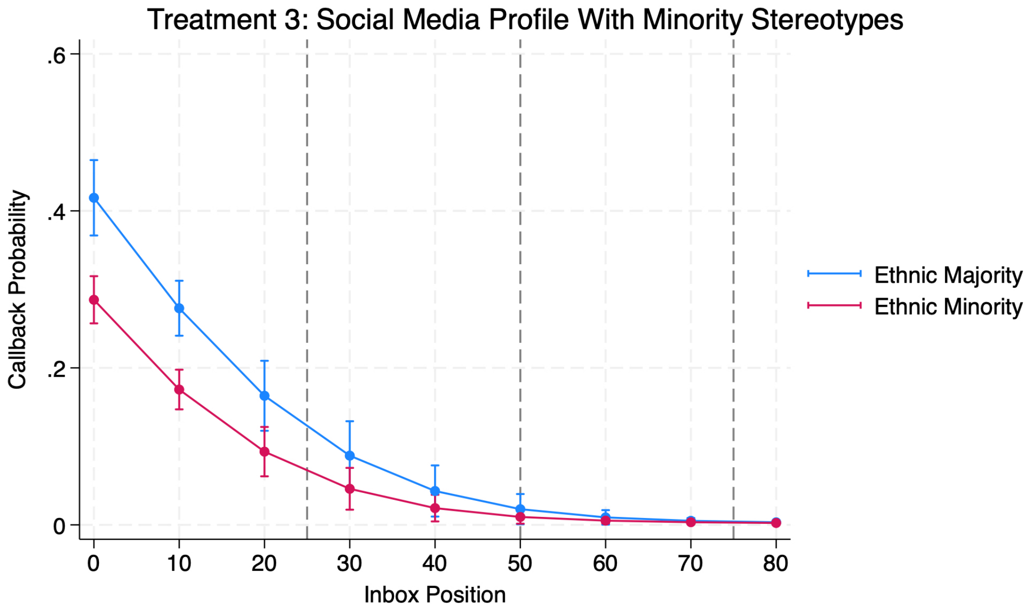


Figure C.6: Predicted Callback Probability by Inbox Position and Ethnicity (Treatment 3)

Table C.8: The Effects of Higher Salience on Callback Probability (Split Samples for Experimental Waves)

Callback	(1) First Wave	(2) Second Wave	(3) First Wave	(4) Second Wave	(5) First Wave	(6) Second Wave
High Salience	0.0339*	0.0333**	0.0310	0.0380***	0.0705*	-0.0236
	(0.0175)	(0.0140)	(0.0204)	(0.0142)	(0.0421)	(0.0288)
Ethnic Minority	-	-	-0.161***	-0.138***	-0.159***	-0.140***
			(0.0246)	(0.0193)	(0.0243)	(0.0194)
High Salience × Ethnic Minority	-	-	0.00419	-0.0102	-0.00109	-0.00795
			(0.0387)	(0.0211)	(0.0363)	(0.0215)
Treatment 1: Without SM (Ref.)	-	-	-	-	-	-
Treatment 2: SM Without Minority Stereotypes	-	-	0.0462*	0.0362*	0.0430	-0.00396
			(0.0243)	(0.0186)	(0.0445)	(0.0305)
Treatment 3: SM With Minority Stereotypes	-	-	-0.0373*	-0.0659***	0.0216	-0.118***
			(0.0197)	(0.0214)	(0.0273)	(0.0286)
High Salience × Treatment 1 (Ref.)	-	-	-	-	-	-
High Salience × Treatment 2	-	-	-	-	0.00388	0.0812
					(0.0657)	(0.0503)
High Salience × Treatment 3	-	-	-	-	-0.120**	0.103***
					(0.0585)	(0.0314)
Observations	2,044	2,079	2,044	2,079	2,044	2,079
Pseudo R^2	0.129	0.134	0.162	0.167	0.166	0.169
Week & Wave FE	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes
All Additional Controls	Yes	Yes	Yes	Yes	Yes	Yes
P-value Heteroscedasticity (Wald) Test	0.303	0.617	0.509	0.485	0.461	0.314

Note: The table reports similar specifications as in Table 4.1 using split samples for each wave. The first wave was conducted between October and November 2023, while the second wave took place between September and November 2024. All additional controls include room & shared apartment, roommate, advertiser, ad statistic, geographic, and district level demographic controls (see notes to Table 4.1 for a detailed description. Robust standard errors (in parentheses) are clustered at the city level and robust to heteroscedasticity. The Wald test yields $\chi^2 = 1.06$ ($p = 0.303$) for model 1, $\chi^2 = 0.25$ ($p = 0.617$) for model 2, $\chi^2 = 0.43$ ($p = 0.509$) for model 3, $\chi^2 = 0.49$ ($p = 0.485$) for model 4, $\chi^2 = 0.54$ ($p = 0.461$) for model 5, and $\chi^2 = 1.02$ ($p = 0.314$) for model 6, indicating no significant heteroscedasticity (Heckman, 1998; Neumark, 2012). * p < 0.1, ** p < 0.05, *** p < 0.01.

Table C.9: The Effects of Ad Quality & Information Salience on Callback Probability

Callback	(1)	(2)
High Salience	0.0453	0.0382
	(0.0317)	(0.0350)
Ethnic Minority	-0.115***	-0.127***
	(0.0406)	(0.0353)
High Salience × Ethnic Minority	0.0205	0.0258
	(0.0407)	(0.0367)
Applications per Visit	-0.978***	-0.685***
	(0.175)	(0.172)
Ethnic Minority × Applications per Visit	-0.330	-0.184
	(0.363)	(0.313)
High Salience × Applications per Visit	-0.255	-0.143
	(0.196)	(0.166)
Ethnic Minority × High Salience × Applications per Visit	-0.249	-0.371
	(0.334)	(0.296)
Treatment 1: Without SM (Ref.)	-	-
Treatment 2: SM Without Minority Stereotypes	0.0179	0.0217*
	(0.0168)	(0.0128)
Treatment 3: SM With Minority Stereotypes	-0.0416*	-0.0443*
	(0.0237)	(0.0230)
High Salience × Treatment 1 (Ref.)	-	-
High Salience × Treatment 2	0.0555	0.0484
	(0.0356)	(0.0352)
High Salience × Treatment 3	-0.0135	-0.0109
	(0.0422)	(0.0383)
Observations	4,212	4,130
Pseudo R^2	0.112	0.169
Week & Wave FE	Yes	Yes
City FE	Yes	Yes
All Additional Controls	No	Yes
P-value Heteroscedasticity (Wald) Test	0.523	0.703

Note: The table reports similar specifications as in Table 4.1 including a proxy for ad quality: applications per visit. All additional controls include room & shared apartment, roommate, advertiser, ad statistic, geographic, and district level demographic controls (see notes to Table 4.1 for a detailed description. Robust standard errors (in parentheses) are clustered at the city level and robust to heteroscedasticity. The Wald test yields $\chi^2 = 0.41$ ($p = 0.523$) for model 1 and $\chi^2 = 0.15$ ($p = 0.703$) for model 2, indicating no significant heteroscedasticity (Heckman, 1998; Neumark, 2012). * p < 0.1, ** p < 0.05, *** p < 0.01.

Table C.10: Average Profile Visits and Impressions per Application

	(1) Ethnic Minority	(2) Ethnic Majority	(3) Diff. if Minority	(4) Ratio
Panel A: All Applications				
Profile Visits	0.17	0.25	-0.081***	1.47
	(0.08)	(0.11)	(0.000)	
Impressions	1.32	1.89	-0.563***	1.43
	(0.68)	(0.69)	(0.000)	
Panel B: Without Stereotypes				
Profile Visits	0.16	0.26	-0.104***	1.65
	(0.08)	(0.14)	(0.000)	
Impressions	0.96	1.87	-0.908***	1.95
	(0.60)	(0.70)	(0.000)	
Panel C: With Stereotypes				
Profile Visits	0.18	0.24	-0.057***	1.31
	(0.07)	(0.08)	(0.000)	
Impressions	1.70	1.90	-0.204***	1.12
	(0.55)	(0.68)	(0.000)	

Note: The table reports means and standard deviations (in parentheses) of average profile visits and impressions per application (number of visits and impressions of a given social media profile in a given week divided by the number of applications sent in the same week per condition). Column 3 shows the p-values for a two-sample Wilcoxon rank-sum (Mann-Whitney) test of proportion testing the null hypothesis that the visits and impressions are equal across ethnicities. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

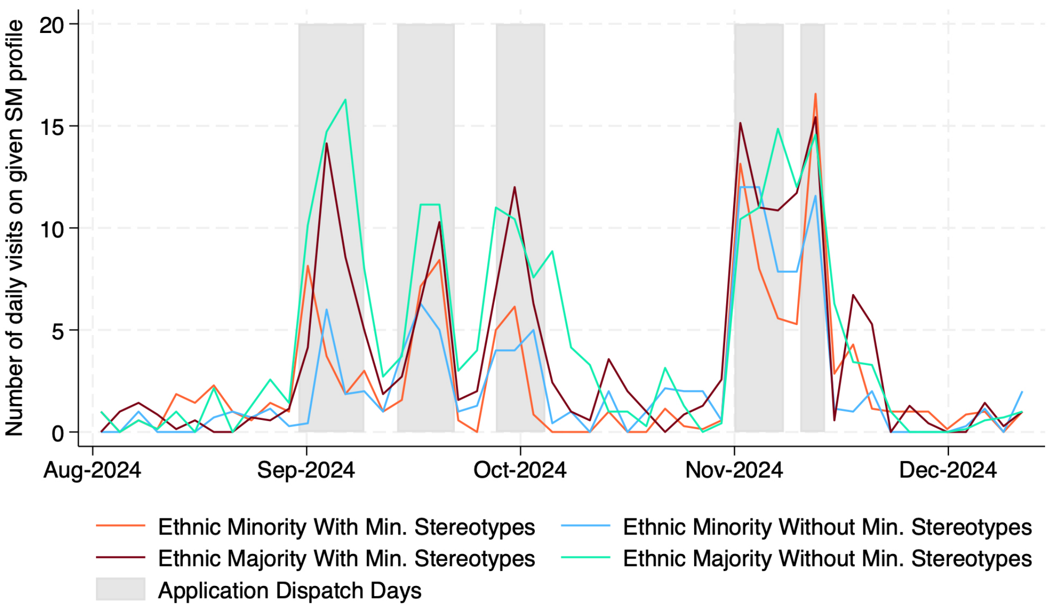


Figure C.7: Daily Profile Visits Over the Course of the Experiment (Second Wave)

C.2 Screenshots

C.2.1 Treatment Notifications

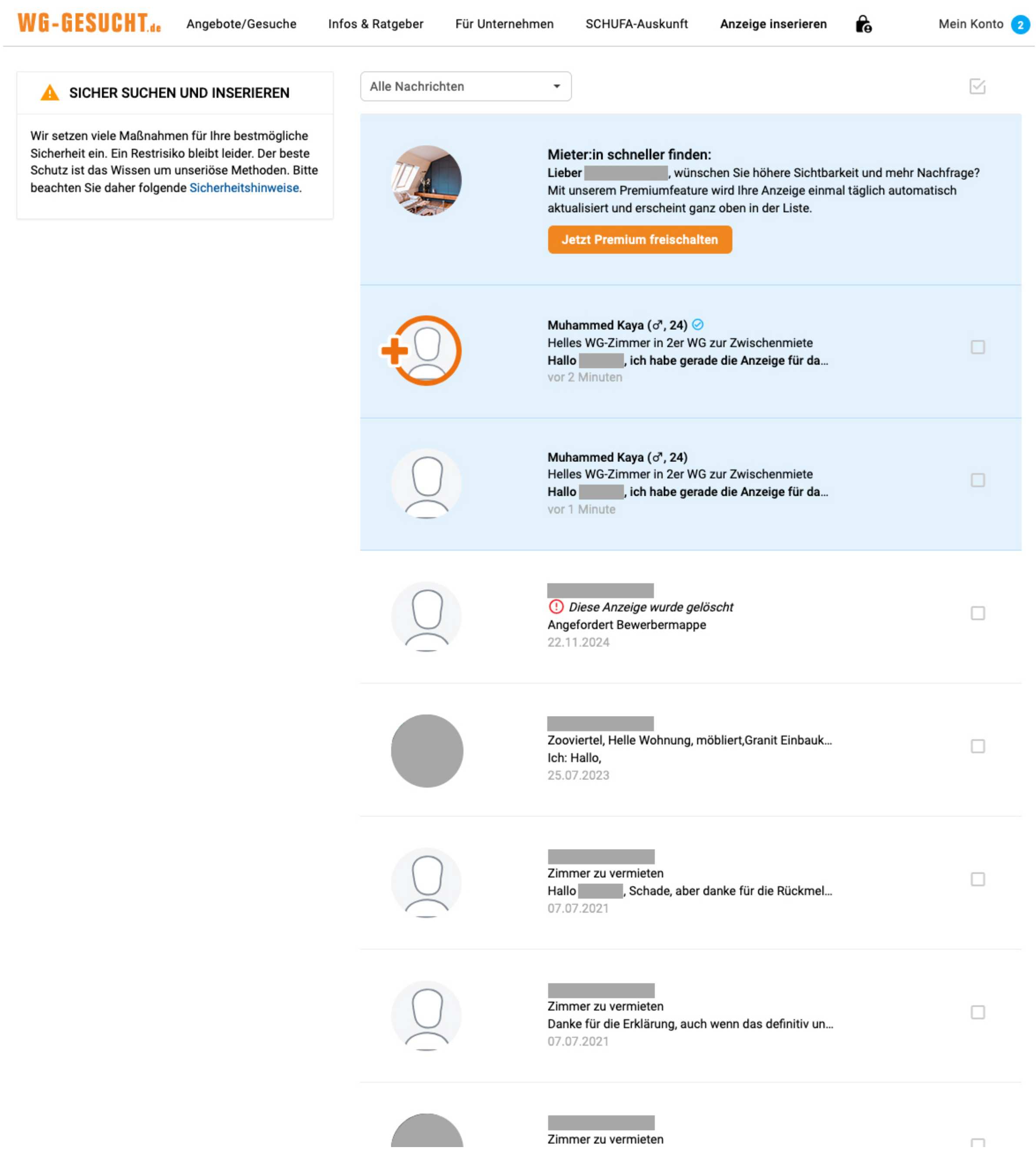


Figure C.8: Screenshot of the Inbox (Example)

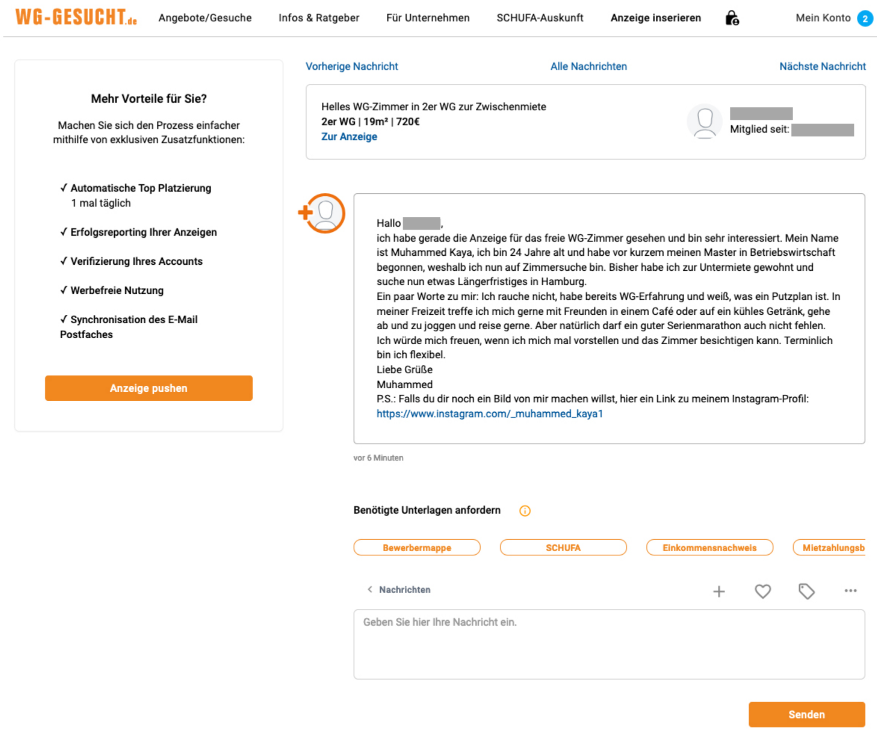


Figure C.9: Screenshot of a Premium Message (Example)

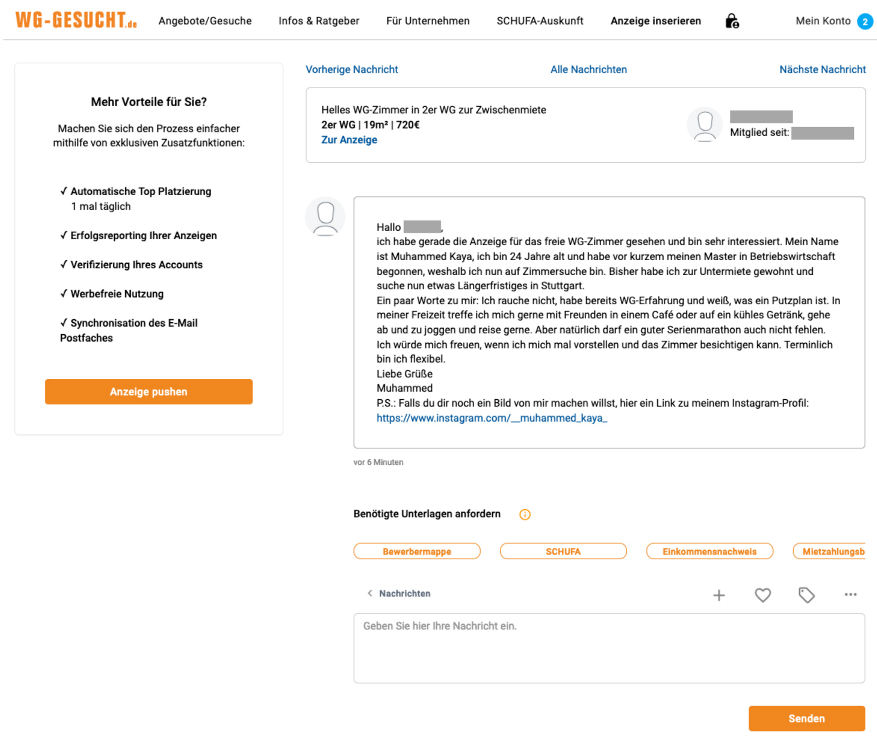


Figure C.10: Screenshot of a Non-Premium Message (Example)

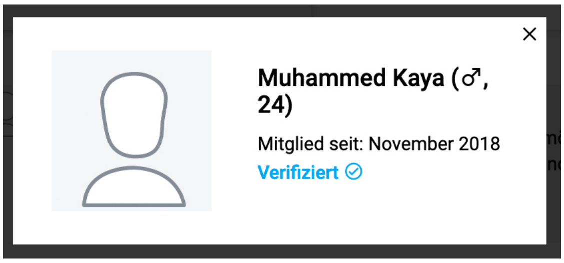


Figure C.11: Screenshot of a User Badge of a Premium User (Example)

C.2.2 Social Media Profiles

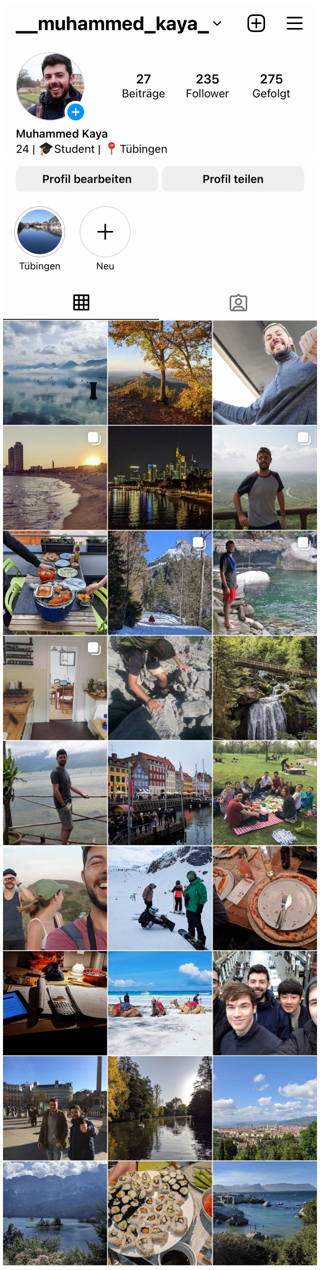


Figure C.12: Screenshot of the Social Media Profile with Turkish-sounding Name Without Visual Minority Stereotypes

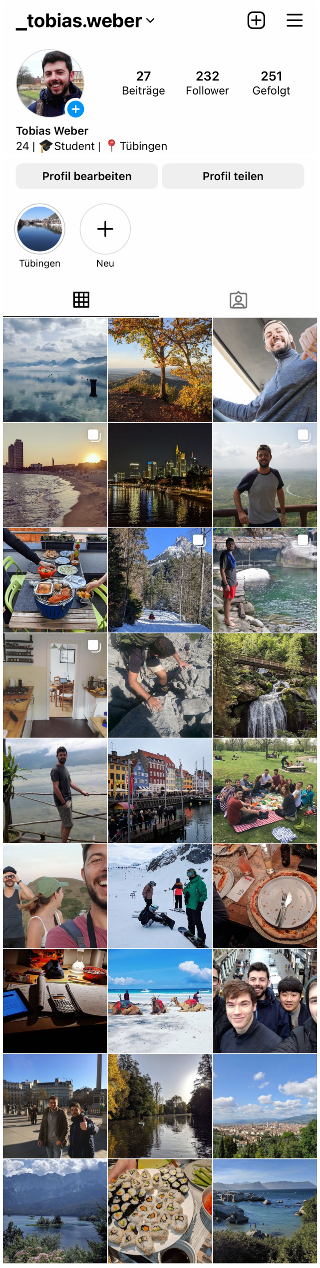


Figure C.13: Screenshot of the Social Media Profile with German-sounding Name Without Visual Minority Stereotypes

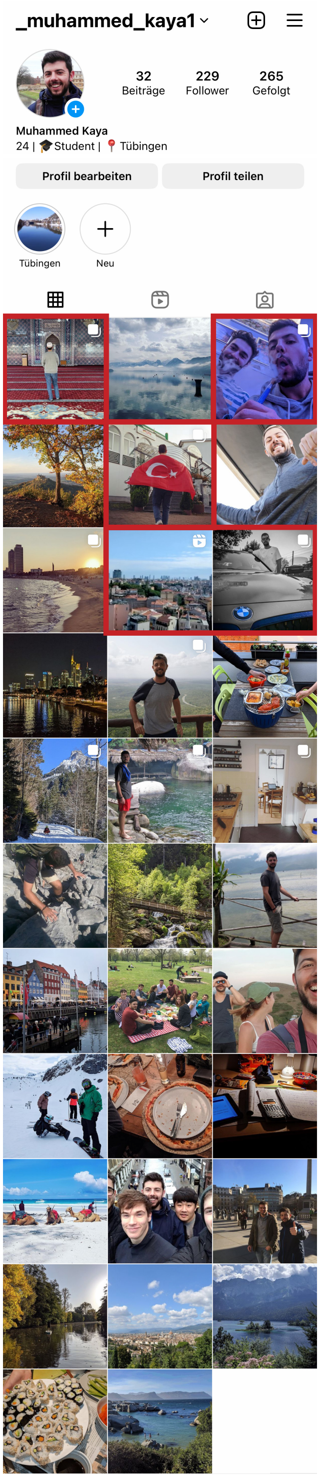


Figure C.14: Screenshot of the Social Media Profile with Turkish-sounding Name With Visual Minority Stereotypes (Red Bordered)

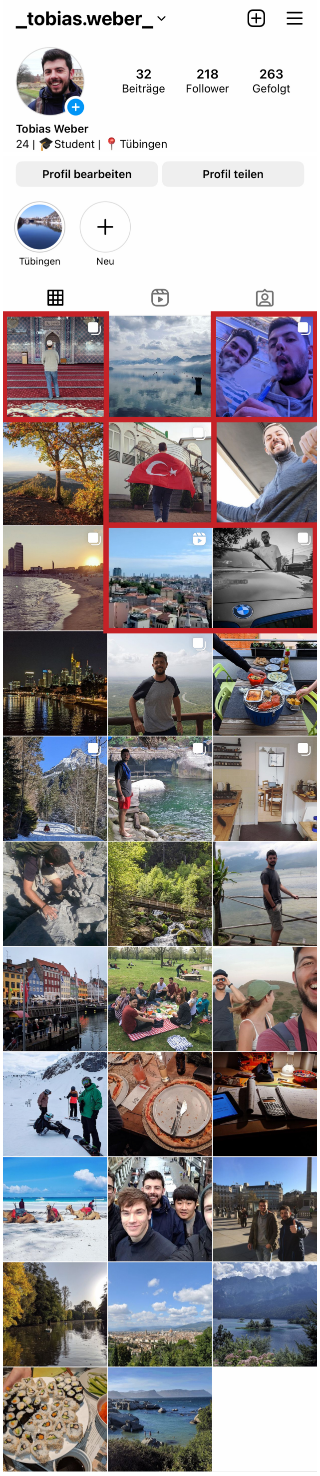


Figure C.15: Screenshot of the Social Media Profile with German-sounding Name With Visual Minority Stereotypes (Red Bordered)

C.3 Premium Subscription Plans of Online Rental Platforms

The following Table C.11 presents information on the premium subscription plans of the most relevant apartment and room rental platforms in the selected country. Given the limited availability of web analytics data (e.g. unique visitors or similar) for a significant proportion of these platforms, I focus on platforms that have been identified by various internet and news sites as being among the most important in the respective countries for the purpose of comparison. I do not include platforms that require a country-specific mobile phone number to register before seeing potential subscription plans.

Columns 2 – 6 report variables for potential tenants and roommates (demand side) while column 7 reports whether premium subscriptions are available for landlords (supply side). Note that special subscriptions or fees (often customized) for professional landlords (large rental agencies and real estate agents that post multiple ads) are not considered as premium subscriptions on the supply side and therefore excluded.

Regarding the demand side, if any kind of premium subscription plan is available for using the site, contacting advertisers, etc., it is shown in column 2. Average monthly subscription/premium fees are indicated in column 3. Fees typically vary with the length of the subscription period or the services offered. Column 4 indicates whether messages from premium users are prioritized in the advertiser's inbox. Column 5 indicates whether certain ads are "premium exclusive" in the sense that they can only be applied to if the prospective tenant or roommate has a premium subscription. Column 6 indicates whether additional services are offered for a fee, such as credit or background checks on the applicant (either included in the subscription or additional one-time payment).

Table C.11: Premium Features of Real Estate Platforms of Selected Countries

	Demand Side					Supply Side
Platform	**Premium Subscription**	**Average Monthly Fee**	**Inbox Prioritization**	**Premium Exclusive Ads**	**Extra Services**	**Premium Subscription**
Germany						
ImmoScout24	Yes	21.49 EUR	Yes	Yes	Yes	Yes
Immowelt	No	.	No	No	No	No
Immonet	No	.	No	No	No	No
WG-gesucht	Yes	17.40 EUR	Yes	No	Yes	No
Wohnungsboerse	No	.	No	No	Yes	No
MeineStadt	No	.	No	No	No	No
Kleinanzeigen	No	.	No	No	No	Yes
WG Börse	No	.	No	No	No	No
wg-liste	No	.	No	No	Yes	No
Immobilo	No	.	No	No	Yes	No
United States						
Zillow	Yes	35.00 USD	No	No	Yes	Yes
Realtor	No	.	No	No	No	Yes
Apartments	Yes	29.00 USD	No	No	Yes	Yes
Trulia (a Zillow brand)	No	.	No	No	No	No
Roomies	No	.	No	No	No	No
Spare Room	Yes	32.19 USD	No	Yes	No	Yes
Roommates	No	.	No	No	No	Yes
Craigslist	No	.	No	No	No	No
United Kingdom						
Rightmove	No	.	No	No	No	No
Zoopla	No	.	No	No	No	No
On the Market	No	.	No	No	Yes	No
Prime Location	No	.	No	No	No	No
Open Rent	No	.	No	No	No	Yes
Spare Room UK	Yes	38.97 GBP	No	Yes	No	Yes
roomies UK	No	.	No	No	No	No
ideal flatmate	No	.	No	No	No	Yes
Roomgo	Yes	31.48 GBP	No	Yes	Yes	Yes
France						
Le Bon Coin	No	.	No	No	Yes	No
Se Loger	No	.	No	No	No	No
Particulier à Particulier	No	.	No	No	Yes	Yes
Avendrealouer	No	.	No	No	No	No
Bienici	No	.	No	No	No	No
La Carte des Colocs	No	.	No	No	No	No
Appartager	Yes	39.98 EUR	No	Yes	Yes	Yes

Continued on next page

Continued from previous page

Platform	Demand Side Premium Subscription	Average Monthly Fee	Inbox Prioritization	Premium Exclusive Ads	Extra Services	Supply Side Premium Subscription
Colocatère	No	.	No	No	No	No
Australia						
realestate.com.au	No	.	No	No	No	No
domain.com.au	No	.	No	No	No	No
allhomes.com.au	No	.	No	No	No	No
flatmates.com.au	Yes	44.98 AUD	No	Yes	Yes	Yes
flatmatesfinders.com.au	No[1]	.	No	No	Yes	No
roommates.com.au	Yes	18.33 AUD	No	No	Yes	Yes
roomgo	Yes	38.48 AUD	No	Yes	No	Yes

Note: The data was collected in January 2025. [1]The platform is free to use for two months but requires a paid subscription afterwards.

Appendix D

Appendix Chapter 5

D.1 Experiment 1: Tables & Figures

Table D.1: Experiment 1: Summary Statistics of Friend Suggestions

		Conscientiousness Mean/SD	**Agreeableness/ Emotional Stability** Mean/SD	Difference
Private	(1/0)	0.653 (0.476)	0.640 (0.480)	-0.013
Posts	(#)	27.574 (58.913)	44.098 (99.587)	16.524**
Followers/subscribers	(#)	466.359 (487.440)	916.810 (7909.810)	450.451
Following/subscriptions	(#)	499.878 (345.310)	555.826 (398.754)	55.948*
Mutual connections	(#)	4.896 (4.242)	3.464 (2.826)	-1.432***
New to Instagram	(1/0)	0.006 (0.077)	0.004 (0.063)	-0.002
Bio	(1/0)	0.608 (0.489)	0.726 (0.446)	0.118***
Emojis (in bio)	(#)	1.076 (1.873)	1.142 (1.964)	0.066
Length of bio	(#)	52.587 (49.441)	48.570 (45.956)	-4.017
Threads account	(1/0)	0.090 (0.286)	0.052 (0.222)	-0.038*
Story highlights (if public)	(#)	1.203 (2.951)	1.420 (3.465)	0.217
Account age	(months)	78.259 (40.157)	91.677 (39.501)	13.418***
Former usernames	(#)	1.108 (2.339)	1.032 (2.262)	-0.076
Female	(1/0)	0.440 (0.497)	0.514 (0.500)	0.074*
Business (if public)	(1/0)	0.026 (0.159)	0.014 (0.118)	-0.012
Non-profit (if public)	(1/0)	0.006 (0.077)	0.008 (0.089)	0.002
Active story (if public)	(1/0)	0.056 (0.230)	0.050 (0.218)	-0.006
Locations (in bio)	(#)	0.388 (0.682)	0.534 (0.875)	0.146**
Tübingen (in bio)	(1/0)	0.020 (0.140)	0.054 (0.226)	0.034**
Countries (in bio)	(#)	0.384 (0.614)	0.530 (0.806)	0.146**
Germany (in bio)	(1/0)	0.289 (0.454)	0.310 (0.463)	0.021
Age (in bio)	(years)	26.218 (8.286)	25.229 (4.284)	-0.990
Student (in bio)	(1/0)	0.050 (0.218)	0.118 (0.323)	0.068***
Motto/saying (in bio)	(1/0)	0.100 (0.300)	0.110 (0.313)	0.010
Link (in bio)	(1/0)	0.147 (0.355)	0.151 (0.358)	0.003
Obs.		502	500	

Note: The table reports summary statistics for the conscientiousness and agreeableness/emotional stability conditions. Column 5 displays the results of a two-sample t-test with unequal variances. $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table D.2: Experiment 1: Mean Reaction Rates

	High	**Low**	**Difference**	**Ratio**
Accept Rates				
Conscientiousness	42.14	39.05	-3.09	1.08
	[159]	[169]	(0.648)	
Agreeableness/Emotional Stability	45.88	29.14	-16.74**	1.57
	[170]	[151]	(0.003)	
Re-Follow Rates				
Conscientiousness	18.37	23.35	4.98	0.79
	[245]	[257]	(0.207)	
Agreeableness/Emotional Stability	19.92	7.87	-12.05***	2.53
	[246]	[254]	(0.000)	
Rejection Rates				
Conscientiousness	17.96	21.01	3.05	0.85
	[245]	[257]	(0.454)	
Agreeableness/Emotional Stability	13.01	22.44	9.43**	0.58
	[246]	[254]	(0.008)	

Note: The table reports mean reaction rates for different experimental conditions. The numbers in brackets in each cell present the number of requests sent for the given sub-sample and high/low conditions. Column 5 shows the p-value for a two-sample Wilcoxon rank-sum (Mann-Whitney) test of proportion testing the null hypothesis that the reaction rates are equal across high/low conditions. $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Figure D.1: Experiment 1: Mean Rejection Rates

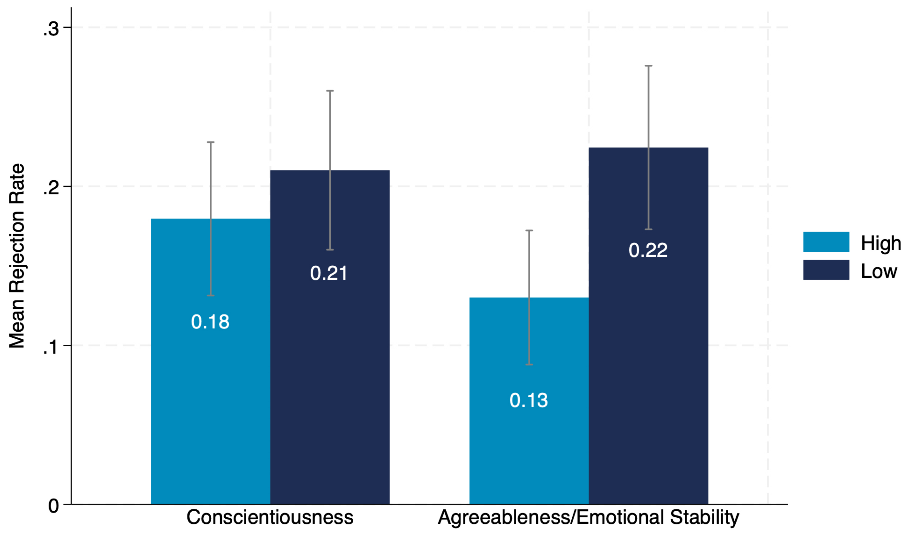


Table D.3: Experiment 1: The Effect of User's Biographic Information on Reaction Rates (Average Marginal Effects)

	(1)	(2)	(3)	(4)
			Agreeableness/	
	Conscientiousness		Emotional Stability	
	Accept	Re-Follow	Accept	Re-Follow
Treatment: Low (Ref.)	-	-	-	-
Treatment: High	-0.014	0.046	0.077	0.026
	(0.056)	(0.035)	(0.054)	(0.034)
Posts	-0.000	-0.000	-0.000	-0.000
	(0.001)	(0.000)	(0.001)	(0.000)
Followers/subscribers	-0.000	-0.000†	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Following/subscriptions	0.000	0.000**	0.000	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)
Mutual connections	0.005	0.004	0.007	0.006†
	(0.007)	(0.003)	(0.007)	(0.003)
Account age (months)	0.001†	0.001	0.001†	0.000
	(0.001)	(0.000)	(0.001)	(0.000)
Female	-0.241***	-0.168***	-0.243***	-0.168***
	(0.044)	(0.030)	(0.044)	(0.030)
Private	-	0.053	-	0.052
		(0.033)		(0.033)
Length of bio	-0.000	-0.000	-0.000	-0.000
	(0.001)	(0.000)	(0.001)	(0.000)
Tübingen in bio	-0.040	0.059	-0.043	0.048
	(0.114)	(0.064)	(0.115)	(0.065)
No. of countries in bio	-0.020	-0.006	-0.019	-0.008
	(0.055)	(0.035)	(0.054)	(0.035)
Student status in bio	0.042	0.034	0.038	0.037
	(0.079)	(0.048)	(0.079)	(0.048)
Motto/saying in bio	-0.192*	0.038	-0.188*	0.041
	(0.077)	(0.042)	(0.077)	(0.042)
Link in bio	0.008	0.057†	0.014	0.061†
	(0.061)	(0.034)	(0.061)	(0.034)
Obs.	404	632	404	632
Additional Controls	Yes	Yes	Yes	Yes
Additional Bio Controls	Yes	Yes	Yes	Yes
Pseudo R^2	0.107	0.116	0.111	0.114
P-value Heteroscedasticity (Wald) Test	0.555	0.205	0.357	0.129

Note: The table reports the same specifications as in Table 5.1 including control variables on a user's bio. The "bio" section can be found at the top of a user's profile page, where users may add a personal description of up to 150 characters. Additional bio control variables include a dummy variable that equals one if the user indicates residence in Germany and the total number of locations given in the bio section. Robust standard errors in parentheses. For a description on additional controls, see the notes to Table 5.1. Robust standard errors (in parentheses) are also robust to heteroscedasticity. The Wald tests yield $\chi^2 = 0.35$ ($p = 0.555$) for model 1, $\chi^2 = 1.61$ ($p = 0.205$) for model 2, $\chi^2 = 0.85$ ($p = 0.357$) for model 3, and $\chi^2 = 2.31$ ($p = 0.129$) for model 4, indicating no significant heteroscedasticity (Heckman, 1998; Neumark, 2012). $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table D.4: Experiment 1: Weekly Profile Visits (Unique Weekly Visits)

	Mean (SD)	Min	Max
High Conscientiousness	1.555 (0.139)	1.462	1.781
Low Conscientiousness	2.074 (0.233)	1.741	2.688
Diff.	0.519*** (0.000)		
High Agreeableness/Emotional Stability	1.951 (0.194)	1.375	2.083
Low Agreeableness/Emotional Stability	1.583 (0.057)	1.429	1.641
Diff.	-0.369*** (0.000)		
High	1.754 (0.260)	1.375	2.083
Low	1.829 (0.299)	1.429	2.688
Diff.	0.076*** (0.000)		
Conscientiousness	1.821 (0.323)	1.462	2.688
Agreeableness/Emotional Stability	1.764 (0.233)	1.375	2.083
Diff.	-0.057** (0.002)		
Total	1.792 (0.283)	1.375	2.688

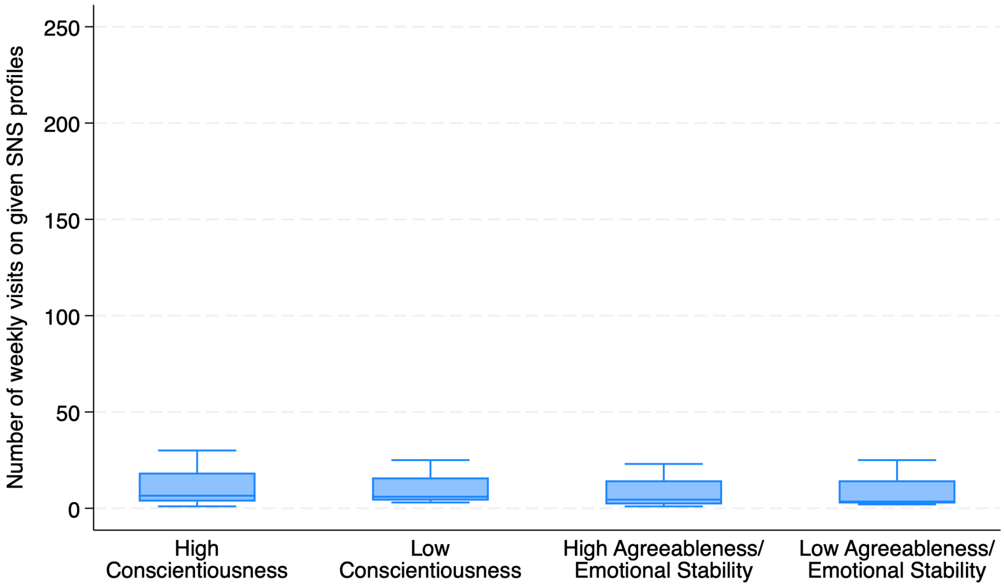
Note: The table reports summary statistics for different experimental conditions on weekly profile visits (unique visits) per request. Profile visits are defined as the number of times a profile was accessed (i.e., unique visits) by a user who viewed a post or story (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]). The values in parentheses for the differences in means report the p-values of two-sample t-tests with equal variances. $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table D.5: Experiment 1: Daily Profile Visits (Unique Daily Visits; Approximated)

	Mean (SD)	Min	Max
High Conscientiousness	1.914 (1.248)	1.206	8.600
Low Conscientiousness	2.601 (1.383)	1.071	6.000
Diff.	0.687*** (0.000)		
High Agreeableness/Emotional Stability	2.772 (1.339)	1.091	5.969
Low Agreeableness/Emotional Stability	1.919 (1.083)	1.119	7.333
Diff.	-0.853*** (0.000)		
High	2.344 (1.362)	1.091	8.600
Low	2.262 (1.288)	1.071	7.333
Diff.	-0.082 (0.328)		
Conscientiousness	2.266 (1.362)	1.071	8.600
Agreeableness/Emotional Stability	2.339 (1.287)	1.091	7.333
Diff.	0.073 (0.383)		
Total	2.302 (1.3253)	1.071	8.600

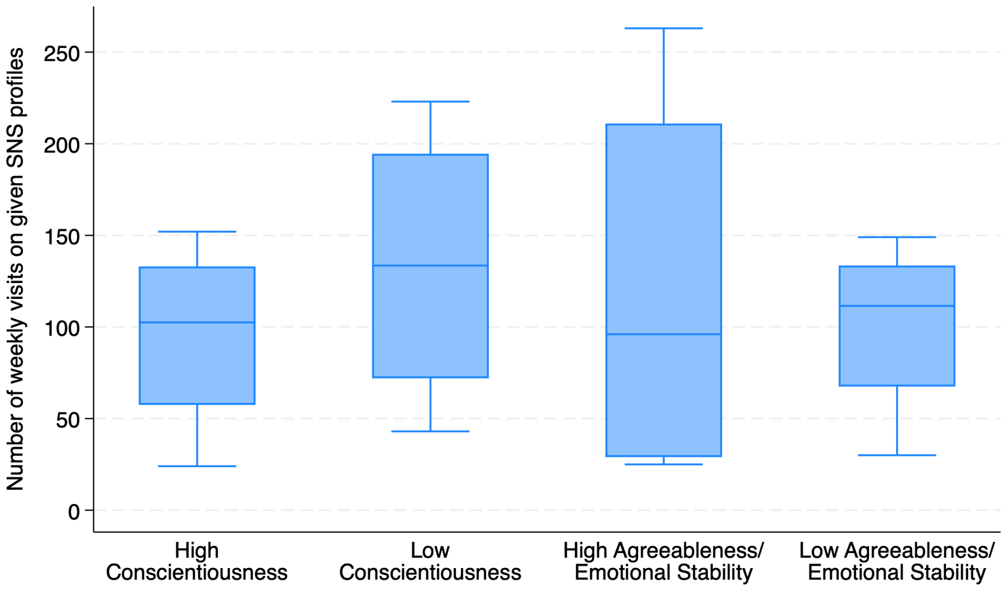
Note: The table reports summary statistics for different experimental conditions on daily profile visits (unique visits) per request. Profile visits are defined as the number of times a profile was accessed (i.e., unique visits) by a user who viewed a post or story (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]). The daily measure here is an approximation, as Instagram provides only limited data on visits (e.g., profile visits from July 3-4 and July 4-5). Therefore, daily profile visits per request is computed as the average of two data points. The values in parentheses for the differences in means report the p-values of two-sample t-tests with equal variances. $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Figure D.2: Experiment 1: Weekly Profile Visits Before the Experiment



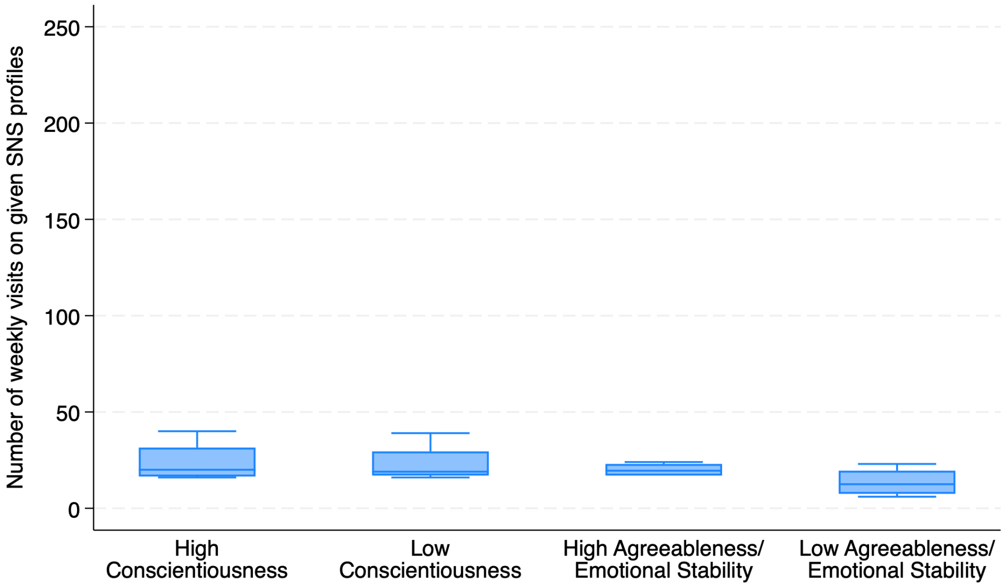
Note: Profile visits are defined as the number of times a profile was accessed (i.e., unique visits) by a user who viewed a post or story (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Figure D.3: Experiment 1: Weekly Profile Visits During the Experiment



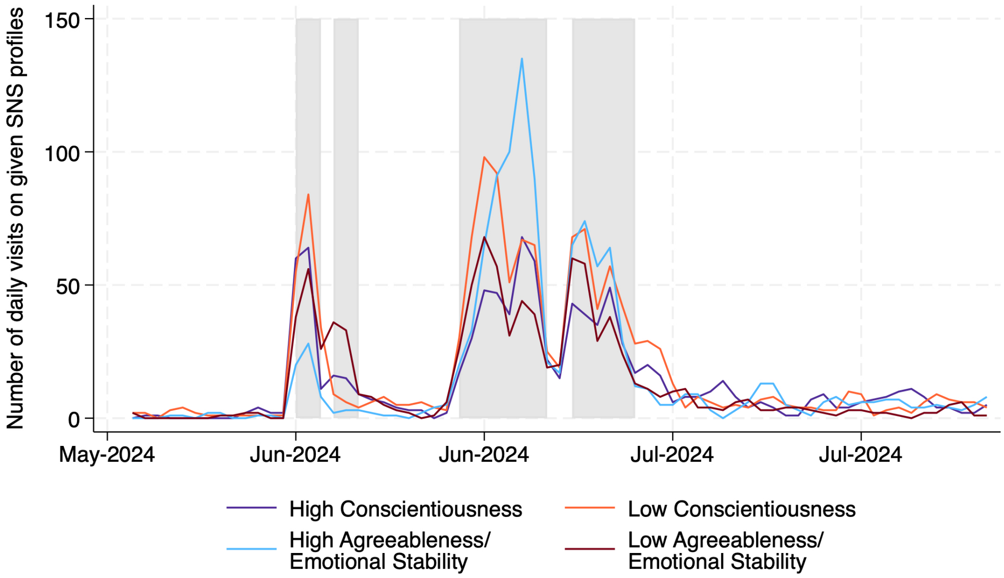
Note: Profile visits are defined as the number of times a profile was accessed (i.e., unique visits) by a user who viewed a post or story (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Figure D.4: Experiment 1: Weekly Profile Visits After the Experiment



Note: Profile visits are defined as the number of times a profile was accessed (i.e., unique visits) by a user who viewed a post or story (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Figure D.5: Experiment 1: Daily Profile Visits Over the Course of the Experiment



Note: Profile visits are defined as the number of times a profile was accessed (i.e., unique visits) by a user who viewed a post or story (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).
The highlighted sections indicate the days on which friend requests were sent.

D.1.1 Robustness Checks, Additional Analyses & Figures

Table D.6: Experiment 1: Robustness Check: Changes in Mutual Connections

	(1)	(2)	(3)	(4)
	Conscientiousness		Agreeableness/ Emotional Stability	
	Accept	Re-Follow	Accept	Re-Follow
Treatment: Low (Ref.)	-	-	-	-
Treatment: High	-0.030	-0.097*	0.196***	0.138**
	(0.067)	(0.050)	(0.053)	(0.051)
Mutual connections	0.003	0.008	0.007	0.011*
	(0.007)	(0.005)	(0.007)	(0.005)
Changes in mutual connections	-0.038†	-0.065**	0.013	-0.008
	(0.022)	(0.021)	(0.015)	(0.012)
Low Consc. × Chg. in mutual con.	0.037	0.048†	-	-
	(0.033)	(0.026)		
High Consc. × Chg. in mutual con.	0.046†	0.056*	-	-
	(0.028)	(0.026)		
Low Agreea. × Chg. in mutual con.	-	-	-0.048†	-0.068*
			(0.025)	(0.027)
High Agreea. × Chg. in mutual con.	-	-	-0.052	-0.038
			(0.055)	(0.044)
Obs.	430	478	430	478
All Controls	Yes	Yes	Yes	Yes
Pseudo R^2	0.0880	-	0.109	0.218
P-value Heteroscedasticity (Wald) Test	0.139	0.042	0.717	0.737

Note: The table reports average marginal effects computed from different probit models with accept (columns 1 and 3) and re-follow (column 4) as the dependent variable and a heteroskedastic probit model with re-follow as the dependent variable (column 2). All specifications include standard and additional control variables (see Table 5.1), the number of changes in mutual connections (between request and reaction) as an additional robustness check to address a potential source of endogeneity, and interaction effects of the treatment dummies and the number of changes in mutual connections. Robust standard errors (in parentheses) are robust to heteroscedasticity. The Wald tests yield $\chi^2 = 2.19$ ($p = 0.139$) for model 1, $\chi^2 = 4.13$ ($p = 0.042$) for model 2, $\chi^2 = 0.13$ ($p = 0.717$) for model 3, and $\chi^2 = 0.11$ ($p = 0.737$) for model 4, indicating no significant heteroscedasticity (Heckman, 1998; Neumark, 2012) for models 1, 3, and 4. The remaining model is computed as heteroskedastic probit model to account for heteroscedasticity (see Section 5.5.3). $^\dagger p < 0.1$, $^* p < 0.05$, $^{**} p < 0.01$, $^{***} p < 0.001$.

Table D.7: Experiment 1: Robustness Check: Heterogeneities in Friend Suggestions I – Conscientiousness

	(1) Accept	(2) Re-Follow	(3) Accept	(4) Re-Follow	(5) Accept	(6) Re-Follow	(7) Accept	(8) Re-Follow
Treatment: Low (Ref.)	-	-	-	-	-	-	-	-
Treatment: High	0.008	-0.016	-0.062	-0.093†	-0.024	-0.087*	0.050	-0.108*
	(0.053)	(0.039)	(0.085)	(0.048)	(0.069)	(0.042)	(0.077)	(0.043)
Posts	-0.002*	-0.001†						
	(0.001)	(0.000)						
Low Consc. × Posts	0.002†	0.001*						
	(0.001)	(0.000)						
High Consc. × Posts	0.001	-0.002						
	(0.003)	(0.002)						
Following/subscriptions			0.000	0.000				
			(0.000)	(0.000)				
Low Consc. × Following			0.000	0.000***				
			(0.000)	(0.000)				
High Consc. × Following			0.000	0.000†				
			(0.000)	(0.000)				
Mutual connections					0.001	-0.009†		
					(0.009)	(0.005)		
Low Consc. × Mutual connections					-0.001	0.017***		
					(0.009)	(0.005)		
High Consc. × Mutual connections					0.008	0.019*		
					(0.012)	(0.008)		
Bio							0.018	-0.059†
							(0.052)	(0.031)
Low Consc. × Bio							0.018	0.112***
							(0.060)	(0.034)
High Consc. × Bio							-0.046	0.157*
							(0.092)	(0.070)
Obs.	648	965	648	965	648	965	648	965
All Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.0720	-	0.0700	-	0.0690	-	0.0690	-
P-value Heteroscedasticity (Wald) Test	0.425	0.007	0.331	0.000	0.517	0.002	0.509	0.001

Note: The table reports the average marginal effects from different probit models with accept as the dependent variable (columns 1, 3, 5, and 7) and from different heteroskedastic probit models with re-follow as the dependent variable (columns 2, 4, 6, and 8). Each specification includes both standard and additional control variables (see Table 5.1), as well as interaction effects between the significantly different characteristics of suggested friends (see Table D.1) and the treatment conditions. Robust standard errors (in parentheses) are also robust to heteroscedasticity. The Wald tests yield $\chi^2 = 0.64$ ($p = 0.425$) for model 1, $\chi^2 = 7.36$ ($p = 0.007$) for model 2, $\chi^2 = 0.95$ ($p = 0.331$) for model 3, $\chi^2 = 14.02$ ($p = 0.000$) for model 4, $\chi^2 = 0.42$ ($p = 0.517$) for model 5, $\chi^2 = 9.98$ ($p = 0.002$) for model 6, $\chi^2 = 0.44$ ($p = 0.509$) for model 7, and $\chi^2 = 11.03$ ($p = 0.001$) for model 8, indicating no significant heteroscedasticity (Heckman, 1998; Neumark, 2012) for models 1, 3, 5, and 7. The remaining models are computed as heteroskedastic probit models to account for heteroscedasticity (see Section 5.5.3). $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table D.8: Experiment 1: Robustness Check: Heterogeneities in Friend Suggestions I – Agreeableness/Emotional Stability

	(1) Accept	(2) Re-Follow	(3) Accept	(4) Re-Follow	(5) Accept	(6) Re-Follow	(7) Accept	(8) Re-Follow
Treatment: Low (Ref.)	-	-	-	-	-	-	-	-
Treatment: High	0.207***	0.150***	0.070	0.053	0.073	0.019	0.129†	0.143**
	(0.053)	(0.043)	(0.075)	(0.047)	(0.087)	(0.047)	(0.072)	(0.054)
Posts	0.000	0.000						
	(0.001)	(0.000)						
Low Agreea. × Posts	-0.001	-0.001†						
	(0.001)	(0.000)						
High Agreea. × Posts	-0.005**	-0.004**						
	(0.002)	(0.001)						
Following/subscriptions			0.000†	0.000***				
			(0.000)	(0.000)				
Low Agreea. × Following			-0.000**	-0.000***				
			(0.000)	(0.000)				
High Agreea. × Following			-0.000	-0.000				
			(0.000)	(0.000)				
Mutual connections					0.010†	0.009*		
					(0.006)	(0.003)		
Low Agreea. × Mutual connections					-0.011	-0.022***		
					(0.009)	(0.006)		
High Agreea. × Mutual connections					0.011	0.010		
					(0.027)	(0.013)		
Bio							0.046	0.055†
							(0.053)	(0.033)
Low Agreea. × Bio							-0.090	-0.134***
							(0.059)	(0.037)
High Agreea. × Bio							-0.054	-0.139*
							(0.088)	(0.056)
Obs.	648	965	648	965	648	965	648	965
All Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.0890	0.116	0.0870	0.120	0.0790	0.115	0.0790	0.115
P-value Heteroscedasticity (Wald) Test	0.957	0.176	0.575	0.872	0.739	0.783	0.697	0.992

Note: The table reports the average marginal effects from different probit models with accept and re-follow as the dependent variables. Each specification includes both standard and additional control variables (see Table 5.1), as well as interaction effects between the significantly different characteristics of suggested friends (see Table D.1) and the treatment conditions. Robust standard errors (in parentheses) are also robust to heteroscedasticity. The Wald tests yield $\chi^2 = 0.00$ ($p = 0.957$) for model 1, $\chi^2 = 1.83$ ($p = 0.176$) for model 2, $\chi^2 = 0.31$ ($p = 0.575$) for model 3, $\chi^2 = 0.03$ ($p = 0.872$) for model 4, $\chi^2 = 0.11$ ($p = 0.739$) for model 5, $\chi^2 = 0.08$ ($p = 0.783$) for model 6, $\chi^2 = 0.15$ ($p = 0.697$) for model 7, and $\chi^2 = 0.00$ ($p = 0.992$) for model 8, indicating no significant heteroscedasticity (Heckman, 1998; Neumark, 2012). $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table D.9: Experiment 1: Robustness Check: Heterogeneities in Friend Suggestions II – Conscientiousness

	(1) Accept	(2) Re-Follow	(3) Accept	(4) Re-Follow	(5) Accept	(6) Re-Follow	(7) Accept	(8) Re-Follow
Treatment: Low (Ref.)	-	-	-	-	-	-	-	-
Treatment: High	0.005	-0.004	0.137	-0.088	0.013	-0.017	-0.037	-0.072*
	(0.048)	(0.038)	(0.101)	(0.059)	(0.065)	(0.043)	(0.051)	(0.036)
Threads account	0.331*	0.115						
	(0.144)	(0.076)						
Low Consc. × Threads	-	-0.107						
		(0.109)						
High Consc. × Threads	-0.233	-0.258*						
	(0.186)	(0.117)						
Account age			$0.001^{\dagger}$	-0.000				
			(0.001)	(0.000)				
Low Consc. × Account age			-0.000	0.001**				
			(0.000)	(0.000)				
High Consc. × Account age			-0.002	0.001				
			(0.001)	(0.001)				
Female					-0.213***	-0.196***		
					(0.045)	(0.034)		
Low Consc. × Female					-0.062	0.078		
					(0.069)	(0.050)		
High Consc. × Female					-0.013	-0.085		
					(0.087)	(0.074)		
Locations (in bio)							-0.038	-0.026
							(0.030)	(0.023)
Low Consc. × Locations							0.038	0.048
							(0.058)	(0.035)
High Consc. × Locations							0.149*	0.108*
							(0.062)	(0.043)
Obs.	640	965	648	965	648	965	648	965
All Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.0750	-	0.0710	-	0.0700	-	0.0750	-
P-value Heteroscedasticity (Wald) Test	0.589	0.059	0.903	0.001	0.772	0.004	0.376	0.008

Note: The table reports the average marginal effects from different probit models with accept as the dependent variable (columns 1, 3, 5, and 7) and from different heteroskedastic probit models with re-follow as the dependent variable (columns 2, 4, 6, and 8). Each specification includes both standard and additional control variables (see Table 5.1), as well as interaction effects between the significantly different characteristics of suggested friends (see Table D.1) and the treatment conditions. Robust standard errors (in parentheses) are also robust to heteroscedasticity. The Wald tests yield $\chi^2 = 0.29$ ($p = 0.589$) for model 1, $\chi^2 = 3.55$ ($p = 0.059$) for model 2, $\chi^2 = 0.01$ ($p = 0.903$) for model 3, $\chi^2 = 12.05$ ($p = 0.001$) for model 4, $\chi^2 = 0.08$ ($p = 0.772$) for model 5, $\chi^2 = 8.39$ ($p = 0.004$) for model 6, $\chi^2 = 0.78$ ($p = 0.376$) for model 7, and $\chi^2 = 7.09$ ($p = 0.008$) for model 8, indicating no significant heteroscedasticity (Heckman, 1998; Neumark, 2012) for models 1, 3, 5, and 7. The remaining models are computed as heteroskedastic probit models to account for heteroscedasticity (see Section 5.5.3). $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table D.10: Experiment 1: Robustness Check: Heterogeneities in Friend Suggestions II – Agreeableness/Emotional Stability

	(1) Accept	(2) Re-Follow	(3) Accept	(4) Re-Follow	(5) Accept	(6) Re-Follow	(7) Accept	(8) Re-Follow
Treatment: Low (Ref.)	-	-	-	-	-	-	-	-
Treatment: High	0.098*	$0.056^{\dagger}$	0.019	0.019	0.104	0.031	0.145**	0.107**
	(0.045)	(0.032)	(0.117)	(0.075)	(0.064)	(0.036)	(0.050)	(0.037)
Threads account	-0.030	-0.064						
	(0.094)	(0.059)						
Low Agreea. × Threads	-	-						
High Agreea. × Threads	$0.294^{\dagger}$	$0.158^{\dagger}$						
	(0.174)	(0.090)						
Account age			0.000	0.000				
			(0.001)	(0.000)				
Low Agreea. × Account age			-0.001	-0.001***				
			(0.001)	(0.000)				
High Agreea. × Account age			0.001	0.000				
			(0.001)	(0.001)				
Female					-0.219***	-0.178***		
					(0.046)	(0.033)		
Low Agreea. × Female					-0.059	$-0.103^{\dagger}$		
					(0.068)	(0.060)		
High Agreea. × Female					0.001	0.061		
					(0.086)	(0.056)		
Locations (in bio)							$0.073^{\dagger}$	0.049*
							(0.040)	(0.024)
Low Agreea. × Locations							-0.112*	-0.075*
							(0.047)	(0.033)
High Agreea. × Locations							$-0.122^{\dagger}$	-0.099*
							(0.064)	(0.039)
Obs.	648	965	648	965	648	965	648	965
All Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.0800	0.101	0.0790	0.114	0.0770	0.105	0.0840	0.108
P-value Heteroscedasticity (Wald) Test	0.979	0.237	0.723	0.987	0.592	0.911	0.542	0.584

Note: The table reports the average marginal effects from different probit models with accept and re-follow as the dependent variables. Each specification includes both standard and additional control variables (see Table 5.1), as well as interaction effects between the significantly different characteristics of suggested friends (see Table D.1) and the treatment conditions. Robust standard errors (in parentheses) are also robust to heteroscedasticity. The Wald tests yield $\chi^2 = 0.00$ ($p = 0.979$) for model 1, $\chi^2 = 1.40$ ($p = 0.237$) for model 2, $\chi^2 = 0.13$ ($p = 0.723$) for model 3, $\chi^2 = 0.00$ ($p = 0.987$) for model 4, $\chi^2 = 0.29$ ($p = 0.592$) for model 5, $\chi^2 = 0.01$ ($p = 0.911$) for model 6, $\chi^2 = 0.37$ ($p = 0.542$) for model 7, and $\chi^2 = 0.30$ ($p = 0.584$) for model 8, indicating no significant heteroscedasticity (Heckman, 1998; Neumark, 2012). $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table D.11: Experiment 1: Robustness Check: Heterogeneities in Friend Suggestions III – Conscientiousness

	(1) Accept	(2) Re-Follow	(3) Accept	(4) Re-Follow	(5) Accept	(6) Re-Follow
Treatment: Low (Ref.)	-	-	-	-	-	-
Treatment: High	0.019	-0.030	-0.027	-0.046	0.002	-0.045
	(0.047)	(0.035)	(0.052)	(0.039)	(0.049)	(0.035)
Tübingen (in bio)	0.010	0.063				
	(0.118)	(0.076)				
Low Consc. × Tübingen	0.140	-				
	(0.383)					
High Consc. × Tübingen	-	0.154				
		(0.207)				
Countries (in bio)			-0.031	-0.023		
			(0.036)	(0.023)		
Low Consc. × Countries			0.042	0.064†		
			(0.067)	(0.038)		
High Consc. × Countries			0.124†	0.045		
			(0.071)	(0.044)		
Student (in bio)					-0.006	-0.027
					(0.085)	(0.049)
Low Consc. × Student					0.027	-0.012
					(0.231)	(0.120)
High Consc. × Student					0.179	0.180†
					(0.167)	(0.108)
Obs.	646	958	648	965	647	963
All Controls	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.0680	-	0.0730	-	0.0690	-
P-value Heteroscedasticity (Wald) Test	0.467	0.018	0.312	0.006	0.497	0.016

Note: The table reports the average marginal effects from different probit models with accept as the dependent variable (columns 1, 3, and 5) and from different heteroskedastic probit models with re-follow as the dependent variable (columns 2, 4, and 6). Each specification includes both standard and additional control variables (see Table 5.1), as well as interaction effects between the significantly different characteristics of suggested friends (see Table D.1) and the treatment conditions. Robust standard errors (in parentheses) are also robust to heteroscedasticity. The Wald tests yield $\chi^2 = 0.53$ ($p = 0.467$) for model 1, $\chi^2 = 5.57$ ($p = 0.018$) for model 2, $\chi^2 = 1.02$ ($p = 0.312$) for model 3, $\chi^2 = 7.56$ ($p = 0.006$) for model 4, $\chi^2 = 0.46$ ($p = 0.497$) for model 5, and $\chi^2 = 5.80$ ($p = 0.016$) for model 6, indicating no significant heteroscedasticity (Heckman, 1998; Neumark, 2012) for models 1, 3, and 5. The remaining models are computed as heteroskedastic probit models to account for heteroscedasticity (see Section 5.5.3). $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table D.12: Experiment 1: Robustness Check: Heterogeneities in Friend Suggestions III – Agreeableness/Emotional Stability

	(1) Accept	(2) Re-Follow	(3) Accept	(4) Re-Follow	(5) Accept	(6) Re-Follow
Treatment: Low (Ref.)	-	-	-	-	-	-
Treatment: High	0.129**	0.074*	0.115*	0.063†	0.098*	0.056†
	(0.044)	(0.032)	(0.052)	(0.037)	(0.046)	(0.032)
Tübingen (in bio)	0.084	0.059				
	(0.292)	(0.133)				
Low Agreea. × Tübingen	0.021	0.028				
	(0.331)	(0.174)				
High Agreea. × Tübingen	-0.214	-0.049				
	(0.336)	(0.161)				
Countries (in bio)			0.077†	0.028		
			(0.046)	(0.028)		
Low Agreea. × Countries			-0.138**	-0.085*		
			(0.053)	(0.039)		
High Agreea. × Countries			-0.063	-0.018		
			(0.072)	(0.045)		
Student (in bio)					0.166	0.096
					(0.122)	(0.070)
Low Agreea. × Student					-0.326*	-0.247*
					(0.157)	(0.106)
High Agreea. × Student					0.014	0.021
					(0.180)	(0.098)
Obs.	647	963	648	965	647	963
All Controls	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.0770	0.0990	0.0840	0.105	0.0830	0.108
P-value Heteroscedasticity (Wald) Test	0.817	0.192	0.682	0.482	0.732	0.512

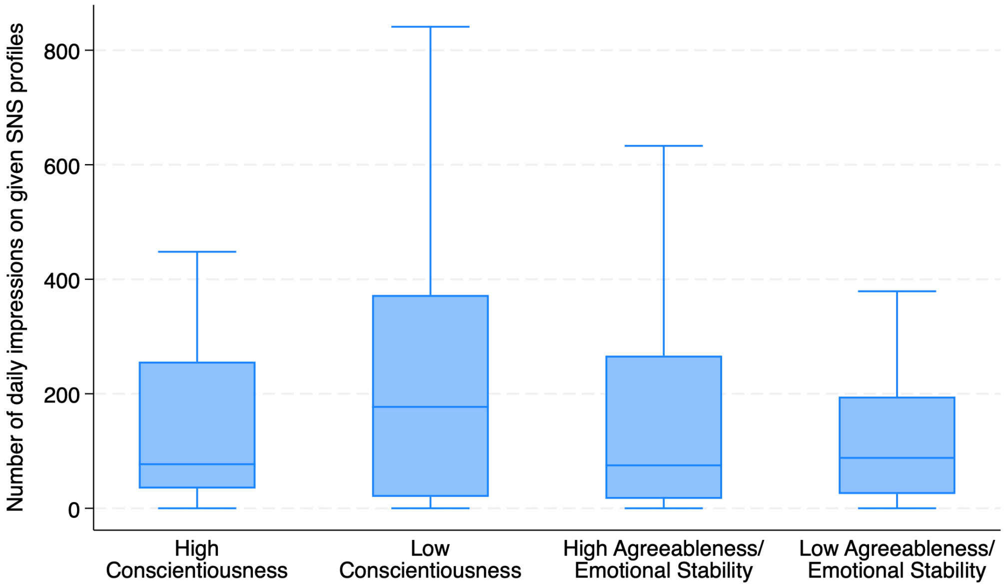
Note: The table reports the average marginal effects from different probit models with accept and re-follow as the dependent variables. Each specification includes both standard and additional control variables (see Table 5.1), as well as interaction effects between the significantly different characteristics of suggested friends (see Table D.1) and the treatment conditions. Robust standard errors (in parentheses) are also robust to heteroscedasticity. The Wald tests yield $\chi^2 = 0.05$ ($p = 0.817$) for model 1, $\chi^2 = 1.70$ ($p = 0.192$) for model 2, $\chi^2 = 0.17$ ($p = 0.682$) for model 3, $\chi^2 = 0.49$ ($p = 0.482$) for model 4, $\chi^2 = 0.12$ ($p = 0.732$) for model 5, and $\chi^2 = 0.43$ ($p = 0.512$) for model 6, indicating no significant heteroscedasticity (Heckman, 1998; Neumark, 2012). $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table D.13: Experiment 1: Determinants of Rejections (Average Marginal Effects)

Rejection	(1) Conscientiousness	(2)	(3) Agreeableness/ Emotional Stability	(4)
Treatment: Low (Ref.)	-	-	-	-
Treatment: High	0.047	0.066**	-0.074***	-0.099***
	(0.030)	(0.032)	(0.026)	(0.026)
Posts		0.000		0.000
		(0.000)		(0.000)
Followers/subscribers		-0.000**		-0.000*
		(0.000)		(0.000)
Following/subscriptions		0.000		0.000
		(0.000)		(0.000)
Mutual connections		-0.002		-0.002
		(0.004)		(0.003)
Account age (months)		-0.000		-0.000
		(0.000)		(0.000)
Bio		0.006		0.002
		(0.029)		(0.030)
Female		0.063**		0.075***
		(0.025)		(0.025)
Private		0.341***		0.339***
		(0.051)		(0.043)
Obs.	1,001	920	1,001	920
All Controls	No	Yes	No	Yes
Pseudo R^2	-	-	0.007	0.165
P-value Heteroscedasticity (Wald) Test	0.003	0.000	0.236	0.218

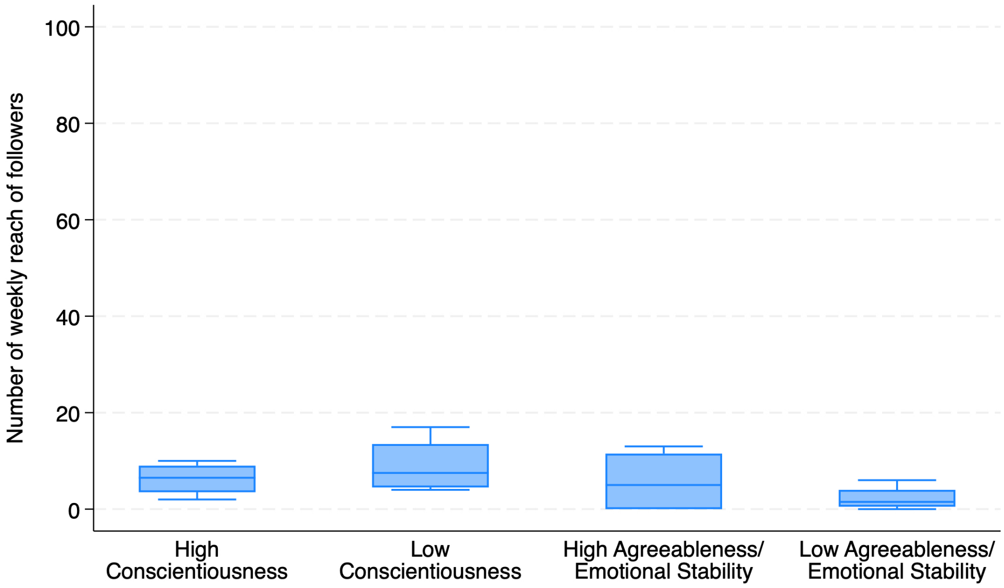
Note: The table reports similar specifications as in Table 5.1 with rejection as the dependent variable. Models in columns 2 and 4 include all control variables. Robust standard errors (in parentheses) are also robust to heteroscedasticity. The Wald tests yield $\chi^2 = 8.94$ ($p = 0.003$) for model 1, $\chi^2 = 13.78$ ($p = 0.000$) for model 2, $\chi^2 = 1.40$ ($p = 0.236$) for model 3, and $\chi^2 = 1.52$ ($p = 0.218$) for model 4, indicating no significant heteroscedasticity (Heckman, 1998; Neumark, 2012) for models 3 and 4. The remaining models are computed as heteroskedastic probit models to account for heteroscedasticity (see Section 5.5.3). $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Figure D.6: Experiment 1: Daily Impressions



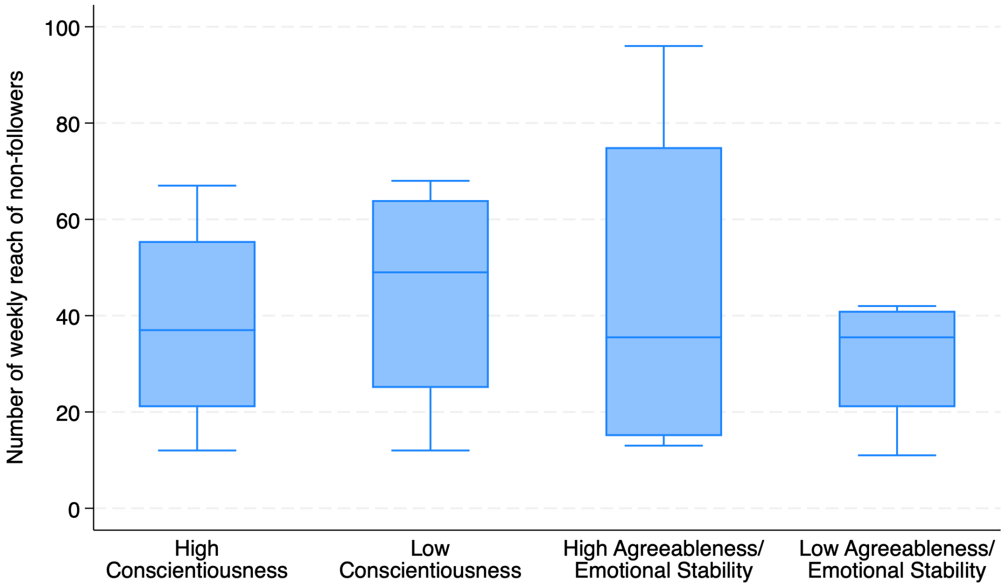
Note: Impressions are defined as the number of times a post or story was on screen (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Figure D.7: Experiment 1: Weekly Reach of Followers



Note: Reach is defined as the number of unique users that have seen posts or stories at least once. It is different from impressions, which may include multiple views by the same user (see: https://help.instagram.com/788388387972460 and https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Figure D.8: Experiment 1: Weekly Reach of Non-Followers



Note: Reach is defined as the number of unique users that have seen posts or stories at least once. It is different from impressions, which may include multiple views by the same user (see: https://help.instagram.com/788388387972460 and https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

D.2 Experiment 2: Tables & Figures

Table D.14: Experiment 2: Summary Statistics

		Mean	**S.D.**	**Obs.**
Experimental Variables				
Treatment 1: No SM	(1/0)	0.196	0.397	2,768
Treatment 2: High Conscientiousness	(1/0)	0.193	0.395	2,768
Treatment 3: Low Conscientiousness	(1/0)	0.205	0.404	2,768
Treatment 4: High Agreeableness/Emotional Stability	(1/0)	0.198	0.398	2,768
Treatment 5: Low Agreeableness/Emotional Stability	(1/0)	0.208	0.406	2,768
Number of applications per treatment and week	(#)	122.420	36.232	2,768
Week 1	(1/0)	0.259	0.438	2,768
Week 2	(1/0)	0.291	0.454	2,768
Week 3	(1/0)	0.138	0.345	2,768
Week 4	(1/0)	0.141	0.348	2,768
Week 5	(1/0)	0.171	0.376	2,768
Aachen	(1/0)	0.041	0.199	2,768
Berlin	(1/0)	0.184	0.388	2,768
Bochum	(1/0)	0.021	0.143	2,768
Darmstadt	(1/0)	0.036	0.187	2,768
Dresden	(1/0)	0.038	0.190	2,768
Düsseldorf	(1/0)	0.035	0.183	2,768
Frankfurt am Main	(1/0)	0.069	0.253	2,768
Gießen	(1/0)	0.043	0.202	2,768
Göttingen	(1/0)	0.039	0.195	2,768
Hamburg	(1/0)	0.090	0.286	2,768
Köln	(1/0)	0.057	0.233	2,768
Leipzig	(1/0)	0.073	0.260	2,768
München	(1/0)	0.107	0.309	2,768
Münster	(1/0)	0.040	0.195	2,768
Stuttgart	(1/0)	0.128	0.334	2,768
Response Variables				
Response	(1/0)	0.529	0.499	2,768
Callback	(1/0)	0.387	0.487	2,768
Rejection	(1/0)	0.059	0.235	2,768
Other response	(1/0)	0.083	0.276	2,768
Callback or other response	(1/0)	0.470	0.499	2,768
Number of days between application and response	(1/0)	1.758	4.285	1,464
Response in same week	(1/0)	0.933	0.250	1,464
Response received outside of platform	(1/0)	0.008	0.086	1,464
Number of characters in response text	(#)	266.430	262.398	1,464
2nd message after initial message	(1/0)	0.012	0.107	1,464
Number of smileys and emojis in response text	(#)	0.005	0.076	2,768
Room & Shared Apartment Characteristics				
Roomsize	(sqm)	17.156	7.219	2,766
Total monthly rent (including utilities)	(€)	552.413	198.185	2,768
Monthly rent (excluding utilities)	(€)	474.033	194.179	2,768
Monthly costs (utilities)	(€)	84.971	60.539	2,274
Other monthly costs	(€)	21.170	32.974	1,121
Deposit	(€)	919.103	521.976	2,371
Clearance Payment	(€)	126.623	258.087	884
Pictures included	(1/0)	0.973	0.161	2,768
Temporary	(1/0)	0.160	0.367	2,768

Continued on next page

Table D.14 – continued from previous page

		Mean	S.D.	Obs.
Availability (if temporary)	(days)	421.648	355.916	443
Online time	(minutes)	104.784	71.435	2,768
Online viewing	(1/0)	0.376	0.484	2,768
Only females wanted	(1/0)	0.223	0.417	2,768
All gender wanted	(1/0)	0.777	0.417	2,768
Age preference for applicants	(1/0)	0.647	0.478	2,768
Number of total roommates	(#)	3.352	1.562	2,768
Apartment size	(sqm)	94.197	42.648	2,071
Smoking permitted	(1/0)	0.420	0.494	2,070
Number of characters of ad description	(#)	2115.817	1409.306	2,768
Applicant is asked to include social media profile	(1/0)	0.089	0.285	2,768
Number of smileys in ad text	(#)	0.408	1.904	2,768
Roommate Characteristics				
Female roommates	(#)	0.981	1.051	2,768
Male roommates	(#)	1.073	1.125	2,768
Diverse roommates	(#)	0.013	0.143	2,768
Average age	(years)	26.813	6.197	2,102
Number of languages spoken by roommates	(#)	2.284	1.042	2,383
Roommates speak German	(1/0)	0.844	0.363	2,768
Roommates speak English	(1/0)	0.707	0.455	2,768
Students	(1/0)	0.646	0.478	2,768
Communally	(1/0)	0.428	0.495	2,768
Non-communally	(1/0)	0.098	0.297	2,768
Males only	(1/0)	0.022	0.148	2,768
Females only	(1/0)	0.146	0.353	2,768
Mixed gender	(1/0)	0.482	0.500	2,768
Young professionals	(1/0)	0.052	0.222	2,768
Employed	(1/0)	0.466	0.499	2,768
Students' hall of residence	(1/0)	0.010	0.102	2,768
Vegetarians/vegans	(1/0)	0.048	0.215	2,768
Cross-generational	(1/0)	0.021	0.143	2,768
Single mother/father	(1/0)	0.008	0.087	2,768
Functional	(1/0)	0.033	0.179	2,768
With children	(1/0)	0.011	0.105	2,768
Fraternity	(1/0)	0.060	0.238	2,768
LGBTQIA roommates	(1/0)	0.086	0.280	2,768
Elderly roommates	(1/0)	0.002	0.047	2,768
Disabled roommates	(1/0)	0.018	0.133	2,768
Shared apartment is/will be newly established	(1/0)	0.064	0.245	2,768
Internationals welcome	(1/0)	0.134	0.341	2,768
Advertiser Characteristics				
Account age	(months)	56.507	45.595	2,764
Profile picture	(1/0)	0.495	0.500	2,768
Age	(years)	35.546	13.585	1,450
Female	(1/0)	0.469	0.499	2,618
Premium account	(1/0)	0.018	0.133	1,453
Ad Statistics				
Favorites	(#)	10.752	15.036	2,768
Views	(#)	128.103	147.085	2,768
Applications sent	(#)	22.185	32.183	2,768
Applications sent per view	(#)	0.162	0.172	2,690
Responses from advertiser	(#)	3.151	7.348	2,768

Continued on next page

Table D.14 – continued from previous page

		Mean	**S.D.**	**Obs.**
Avg. duration of advertiser's response	(hours)	16.791	23.461	1,505
Applications from premium account	(#)	3.231	6.225	2,768
Geographic Variables				
Out of town	(1/0)	0.028	0.164	2,768
Distance to city center	(kilometers)	4.708	3.663	2,762
Distance to district center	(kilometers)	2.659	2.312	2,377
Distance to main train station	(kilometers)	4.808	3.652	2,762
Bars in 500m radius	(#)	8.153	7.503	2,767
Cafes in 500m radius	(#)	8.201	7.501	2,767
University/faculty buildings in 1km radius	(#)	7.482	7.774	2,767
Distance to largest university in city	(kilometers)	6.296	5.561	2,765
Distance to 2nd largest university in city	(kilometers)	5.475	4.195	2,765
Distance to 3rd largest university in city	(kilometers)	6.894	5.781	2,302
Avg. distance to largest two universities in city	(kilometers)	5.885	4.575	2,765

Note: The table presents summary statistics on experimental, response, and geographic variables, as well as room & shared apartment, roommates, and advertiser characteristics.

Table D.15: Experiment 2: Mean Callback Rates (SM-Conditions)

	High	**Low**	**Diff. (p-value)**	**Ratio**
Conscientiousness	40.93	37.32	-3.61	1.09
	[535]	[568]	(0.219)	
Agreeableness/Emotional Stability	44.24	31.30	-12.94***	1.41
	[547]	[575]	(0.000)	

Note: The table reports mean callback rates for the treatment conditions including a social media (SM) profile. The second to last column displays the results of a two-sample Wilcoxon rank-sum test. $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table D.16: Experiment 2: Mean Callback Rates Across Conditions

	Callback Rate	**Obs.**	**Diff.**	**(p-value)**
Reference Treatment (no SM)	40.33	543	–	
High Conscientiousness	40.93	535	0.01	(0.840)
Low Conscientiousness	37.32	568	-3.01	(0.304)
High Agreeableness/Emotional Stability	44.24	547	3.91	(0.191)
Low Agreeableness/Emotional Stability	31.30	575	-9.03**	(0.002)

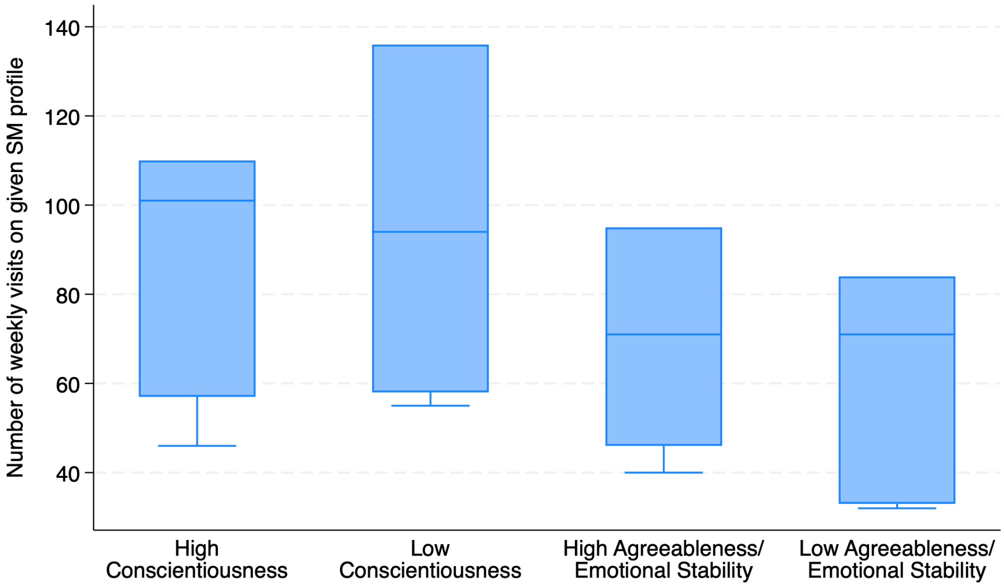
Note: The table reports mean callback rates for all treatment conditions. The last column displays the results of a two-sample test of proportion testing whether the callback rate for each treatment condition is significantly different from the reference treatment. $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table D.17: Experiment 2: Estimated Average Profile Visits, Impressions, and Reach per Application

	(1) High	(2) Low	(3) Diff. (p-value)	(4) Ratio
Panel A: Agreeableness/Emotional Stability				
Profile Visits	0.56	0.48	-0.083***	1.17
	(0.06)	(0.08)	(0.000)	
Impressions	5.11	4.22	-0.890***	1.21
	(1.12)	(0.95)	(0.000)	
Reach of Non-Subscribers	0.35	0.29	-0.061***	1.21
	(0.07)	(0.06)	(0.000)	
Panel B: Conscientiousness				
Profile Visits	0.73	0.72	-0.002*	1.00
	(0.07)	(0.06)	(0.032)	
Impressions	7.31	7.09	-0.217***	1.03
	(0.91)	(0.83)	(0.000)	
Reach of Non-Subscribers	0.40	0.39	-0.000*	1.00
	(0.05)	(0.04)	(0.032)	

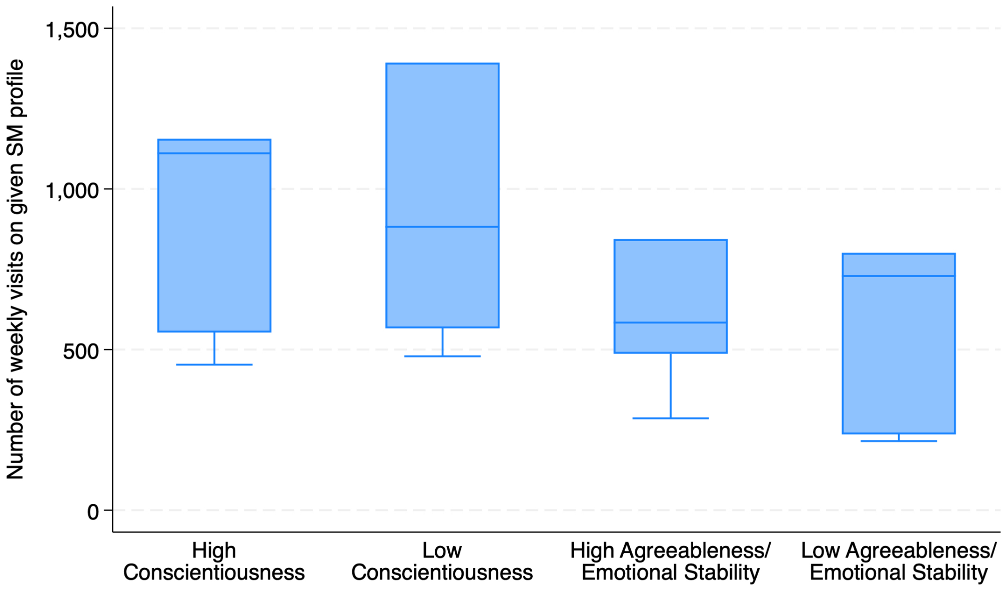
Note: The table reports means and standard deviations (in parentheses) of calculating the estimated average profile visits, impressions, and reach of non-subscribers per application by computing the number of visits, impressions, and reach of a given social media (SM) profile in a given week divided by the number of applications sent per name in the same week for the respective high and low conditions in columns 1 and 2, respectively. Column 3 shows the p-values of a two-sample Wilcoxon rank-sum (Mann-Whitney) test of proportion testing the null hypothesis that the visits, impressions, and reach are equal across high and low conditions. $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Figure D.9: Experiment 2: Weekly Profile Visits



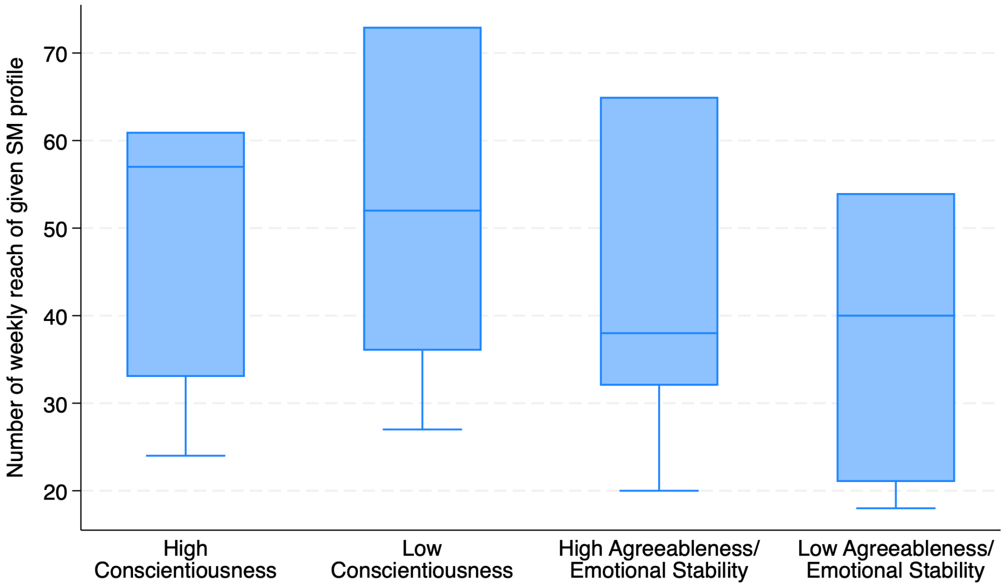
Note: Profile visits are defined as the number of times a profile was accessed (i.e., unique visits) by a user who viewed a post or story (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Figure D.10: Experiment 2: Weekly Impressions



Note: Impressions are defined as the number of times a post or story was on screen (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Figure D.11: Experiment 2: Weekly Reach of Non-Followers



Note: Reach is defined as the number of unique users that have seen posts or stories at least once. It is different from impressions, which may include multiple views by the same user (see: https://help.instagram.com/788388387972460 and https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

D.2.1 Robustness Checks, Additional Analyses & Figures

Table D.18: Experiment 2: Summary Statistics on Social Media Variables

	Conscientiousness				**Agreeableness/Emotional Stability**			
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
	High		**Low**		**High**		**Low**	
Posts	35	0.000	35	0.000	36	0.000	30	0.000
Stories	15	0.000	15	0.000	15	0.000	15	0.000
Followers/subscribers	143	3.147	165	0.000	153	1.177	124	1.212
Following/subscriptions	303	0.435	311	0.000	304	0.819	296	0.000
Two-Day Data								
Visits	28.650	9.004	32.356	16.464	24.236	8.054	20.428	7.069
Impressions	312.654	116.837	324.421	164.021	232.715	95.528	190.856	82.388
Reach of followers	0.413	0.550	0.435	0.496	0.651	0.719	0.190	0.392
Reach of non-followers	16.417	5.430	18.275	9.195	15.344	6.522	13.200	5.143
Weekly Data								
Visits	84.828	24.389	91.546	32.155	68.338	20.383	61.496	19.987
Impressions	868.361	294.661	904.889	345.509	609.572	182.718	560.687	240.049
Reach of followers	1.254	0.667	1.496	0.783	1.263	0.441	0.409	0.492
Reach of non-followers	46.869	14.062	50.266	16.805	42.689	15.757	37.026	13.036
Obs.	535		568		547		575	

Note: The table reports summary statistics of social media metrics over two-day and weekly intervals. While Instagram does not provide daily metrics, it offers data for 48-hour intervals. All measures are derived from unique users. Profile visits are defined as the number of times a profile was accessed (i.e., unique visits) by a user who viewed a post or story. Impressions are defined as the number of times a post or story was on screen (see: https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]). Reach is defined as the number of unique users that have seen posts or stories at least once. It is different from impressions, which may include multiple views by the same user (see: https://help.instagram.com/788388387972460 and https://help.instagram.com/825941707897287 [Retrieved: June 17, 2024]).

Table D.19: Experiment 2: Summary Statistics on District Population Variables

	Mean	S.D.	Obs.
District population 2017	136,564.689	120,039.928	2,689
District population 2018	137,741.715	121,174.239	2,689
District population 2019	138,557.991	121,663.010	2,689
District population 2020	138,487.946	121,646.850	2,689
District population 2021	138,351.460	121,724.113	2,689
District population 2022	140,922.420	124,336.615	2,689
District population 2023	141,784.458	125,286.401	2,689
Population change 2017-2023 (annual growth rate)	0.005	0.009	2,689
Share of female population 2017	0.501	0.019	2,689
Share of female population 2018	0.501	0.018	2,689
Share of female population 2019	0.497	0.038	2,689
Share of female population 2020	0.503	0.016	2,689
Share of female population 2021	0.501	0.018	2,689
Share of female population 2022	0.502	0.017	2,689
Share of female population 2023	0.501	0.017	2,689
District population density 2022	3,552.898	3,107.765	2,576
District population density 2023	3,520.347	3,296.892	2,235
Households	111,938.092	122,589.055	2,576
Apartments	95,877.520	68,822.065	1,847
Average age	41.543	2.131	2,235

Note: The table reports summary statistics for population variables on the district level.

Table D.20: Experiment 2: Effect of Treatment Conditions on Rejections (Average Marginal Effects)

Rejection	(1)	(2)
Treatment: No SM (Ref.)	-	-
Low Agreeableness/Emotional Stability	-0.351*	-0.171†
	(0.179)	(0.102)
Low Conscientiousness	-0.116***	-0.0605†
	(0.0349)	(0.0309)
High Agreeableness/Emotional Stability	-0.232**	-0.125*
	(0.0863)	(0.0551)
High Conscientiousness	-0.0444***	-0.0195
	(0.0133)	(0.0154)
Obs.	2,768	2,530
Week & City FE	Yes	Yes
Room & Shared Apartment Controls	No	Yes
Roommate Controls	No	Yes
Advertiser Controls	No	Yes
Ad Statistic Controls	No	Yes
Geographic & District Level Demographic Controls	No	Yes
P-value Heteroscedasticity (Wald) Test	0.001	0.006

Note: The table reports the average marginal effects computed from two heteroskedastic probit models with rejection as the dependent variable. For a list of control variables, see Table 5.2. The Wald tests yield $\chi^2 = 11.48$ ($p = 0.001$) for model 1 and $\chi^2 = 7.51$ ($p = 0.006$) for model 2, indicating significant heteroscedasticity (Heckman, 1998; Neumark, 2012). Thus, both models are computed as heteroskedastic probit models to account for heteroscedasticity. Standard errors (in parentheses) are clustered at the city level. $^{\dagger}p < 0.1$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

D.3 Experiment 2: Application Text

The following is a non-literal English translation of the application text, which was originally sent in German (variables italicized). The hyperlink to the Instagram profile is automatically converted into a clickable link on the platform for vacant room ads. When accessing the profile from a mobile device with the Instagram application installed (and having logged in), the profile typically opens within the Instagram app, thereby enabling immediate access to all profile details. On a computer, the profile appears in a new browser tab or window, allowing users to view essential information (e.g., username, displayed name, biographical details, and posts) without requiring a login. However, certain information, such as the lists of subscribers and subscriptions, is only accessible when the user is logged in to Instagram.

> Hello *advertiser*,
> I just saw the ad for the vacant room in the shared apartment and I'm very interested. My name is Julia Becker, I am 24 years old and have recently started a Master's in Business and Economics in *city* which is why I'm now looking for a room. I've been subletting so far, but I'm now looking for something longer-term.
> A few words about me: I don't smoke, I've already lived in a shared apartment during my Bachelor's, and I know what a cleaning schedule is. In my spare time, I enjoy meeting up with friends for coffee or a drink, going for a jog, or travel the world. But of course, watching a good TV series from time to time is also a must.
> I would be happy to introduce myself and see the room. I am flexible with dates.
> Best regards
> Julia
> P.S.: If you want to get a picture of me, here's a link to my Instagram profile: https://www.instagram.com/*instagram_profile*

D.4 Screenshots

D.4.1 Social Media Profiles

Figure D.12: Screenshot of the *High Agreeableness/Emotional Stability* Social Media Profile (Mobile)

Figure D.13: Screenshot of the *Low Agreeableness/Emotional Stability* Social Media Profile (Mobile)

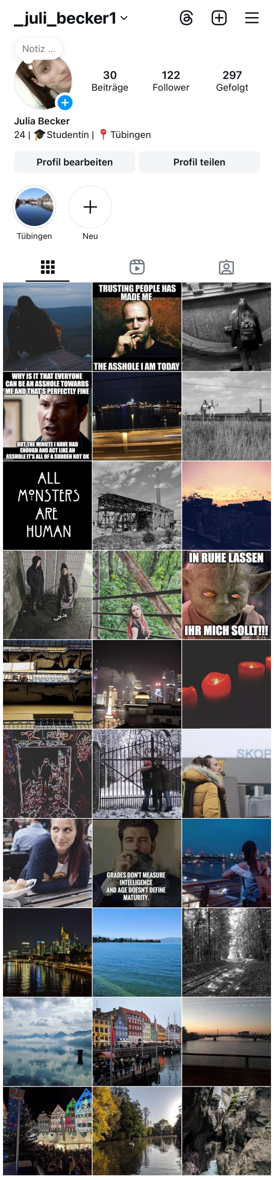


Figure D.14: Screenshot of the *High Conscientiousness* Social Media Profile (Mobile)

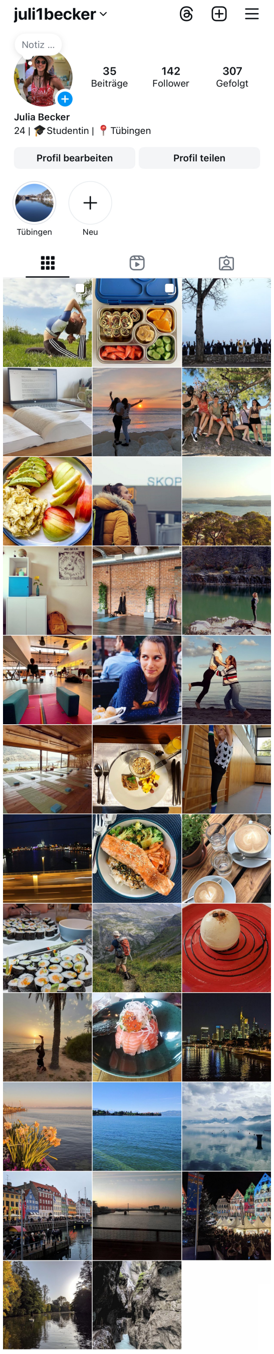
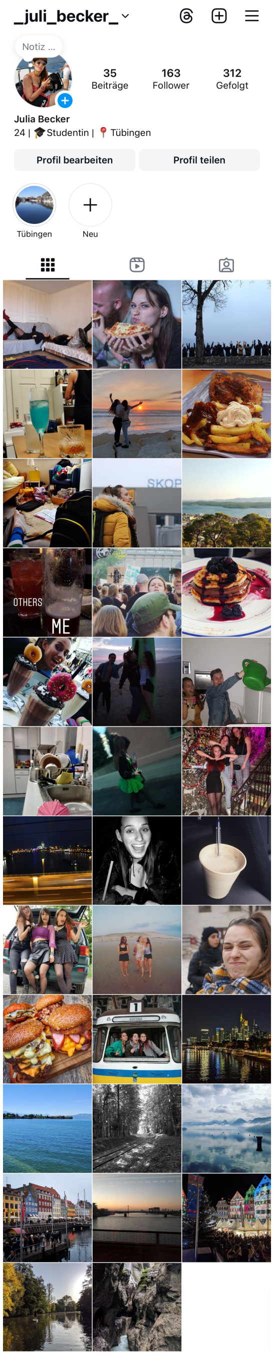


Figure D.15: Screenshot of the *Low Conscientiousness* Social Media Profile (Mobile)



D.4.2 Experiment 1: Friend Suggestions & Treatment Notifications

Figure D.16: Screenshot of Friend Suggestions (Browser)

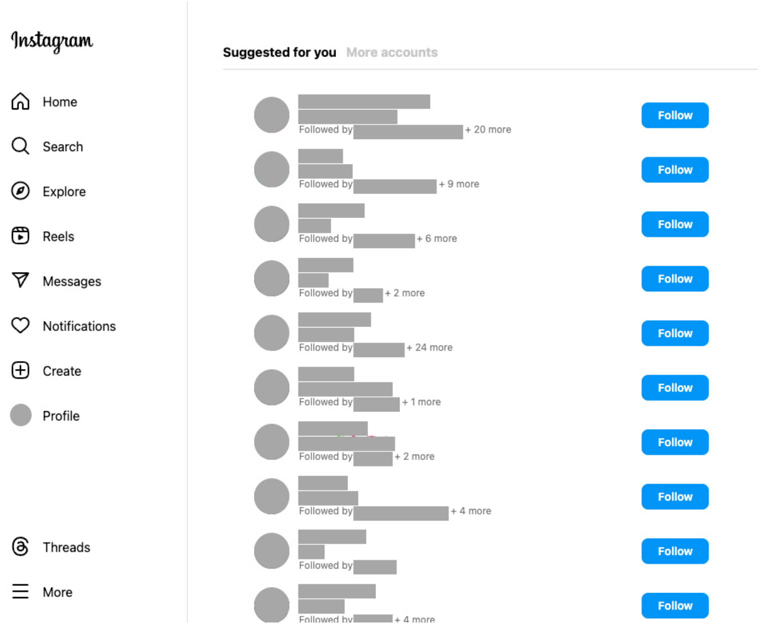
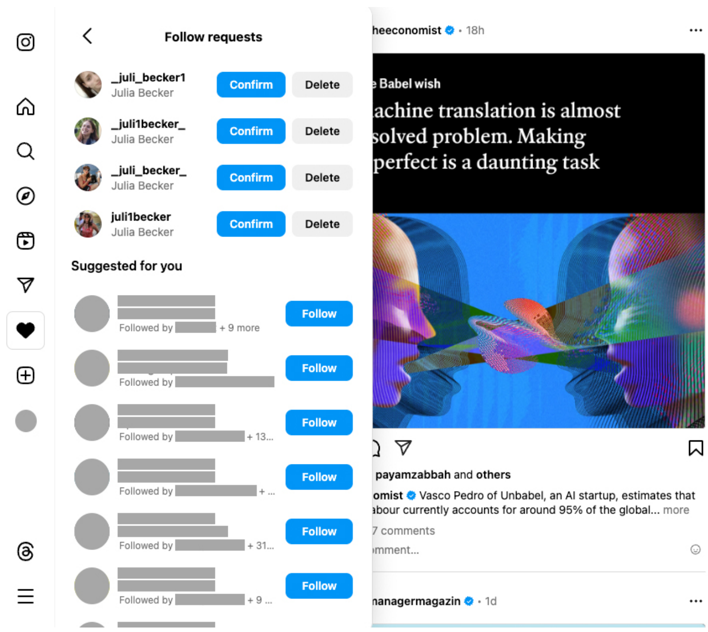
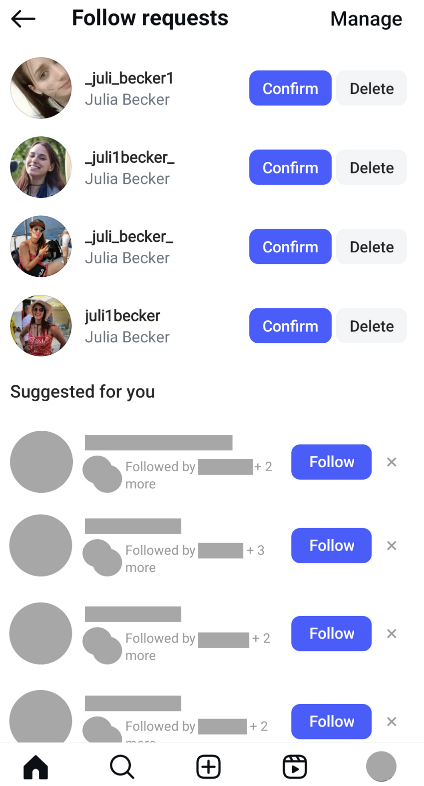


Figure D.17: Example Screenshot of Treatment Notification (Browser)



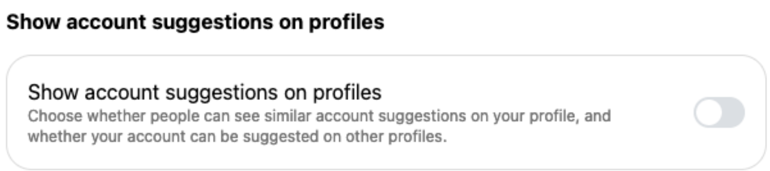
Note: In the experiment, each user received one request only.

Figure D.18: Example Screenshot of Treatment Notification (Mobile)



Note: In the experiment, each user received one request only.

Figure D.19: Screenshot of Profile Setting



D.5 Online Pilot Surveys & Experiments

Overall, we conduct two online pilot experiments to check whether the design of the treatment conditions successfully manipulate relevant beliefs on personality traits of the person shown on the social media profiles (sections D.5.2 and D.5.3). In addition, we conduct three online surveys in order to assess which personality traits should potential roommates posses (Section D.5.1), which personality traits are signaled through the application text (reference treatment) of Experiment 2 (Section D.5.4), and finally, to gain followers for the fictitious profiles to create an initial social network (Section D.5.5).

D.5.1 First Online Survey

We conduct a pilot study with the objective of gaining insights into the demographics, characteristics, and personality traits that potential roommates should possess using the five-factor model (McCrae & Costa, 1987; McCrae & John, 1992). The survey was conducted with students and faculty members at the University of Tübingen in July and August 2020. A total of $n = 787$ respondents participated.

Survey Design Firstly, we ask participants about their current living situation and whether they have prior experience living in shared apartments. Subsequently, participants are asked about their personal preferences in relation to shared apartments and their experiences of searching and selecting new roommates. Those with experience in roommate selection, we additionally inquire more specifically on information acquisition about potential roommates. Secondly, we request subjects to indicate their preferences regarding the personality of a potential roommate applying for a vacant room using the Big Five Inventory-2-S (Rammstedt et al., 2020; Soto & John, 2017). It comprises 30 items, which are presented on a 5-point Likert scale (ranging from strongly disagree to strongly agree).[1]

Thirdly, we inquire about characteristics that a potential roommate or applicant for a vacant room should possess, including experience living in shared apartments, smoking habits, student status, etc.[2] using a 5-point Likert scale. Thereafter, we ask participants about their preferences regarding a potential roommates' sex, origin, and language skills. Finally, participants are asked to provide information regarding their demographic characteristics.

Results We exclude 19 observations from the analysis as these respondents provide inconclusive answers or complete the survey too fast. Most respondents (86.8%) are between the age of 18 and 29 with a mean age of 25.1 years. 76.1% identify as female. 80.9% are students. Furthermore, the vast majority (93.1%) of respondents report to have a German nationality. In contrast to other social media platforms, participants report using Instagram the most.

Regarding the living situation, approximately two thirds of participants indicate to live in a shared apartment.[3] Of these, 88.6% of respondents indicate to live in a shared apartment comprising solely students, while 38.8% state to live communally, i.e., frequently engage in joint activities, such as cooking or watching a movie.

Approximately 61% indicate to have experience in searching and selecting new roommates. Of these, approximately half state that they attempt to obtain additional information about an applicant for a vacant room online or on social media. Additionally, they report receiving an average of 21 to 30 applications for one vacant room.

The results of the second section of the survey, in which participants are asked to indicate their preferences regarding the personality traits of a potential roommate, reveal that conscientiousness and agreeableness exhibit the largest means of all personality traits (see column 1 of Table D.21). Furthermore, both traits exhibit the largest difference from the midpoint of the scale (see column 5 of Table D.21), indicating the strongest preference for these two traits. Conscientiousness and agreeableness are followed by neuroticism/emotional instability with a difference of -1.229 from the midpoint.[4] Extraversion and openness only exhibit a slight difference from the midpoint. Using a one-sample t-test to test the null hypothesis that the mean of the variables is equal to the midpoint of the scale shows that this hypothesis can be rejected for each trait (see column 5 of Table D.21).

In the third section of the survey, respondents are asked to evaluate a series of pre-defined characteristics of a potential roommate, which are deemed essential for cohabitation (see above). Table D.23 presents the results of these items. The respondents indicate that the ability to coexist

[1] The Big Five Inventory-2-S (BFI-2-S) (Rammstedt et al., 2020; Soto & John, 2017) is a psychometric instrument designed to assess the five-factor model of personality, encompassing openness, conscientiousness, extraversion, agreeableness, and neuroticism (emotional instability). This inventory is a brief version of the BFI-2 by Soto and John (2017) and exhibits high reliability and validity (Rammstedt et al., 2020). See Figure D.20 for a screenshot of the personality task.

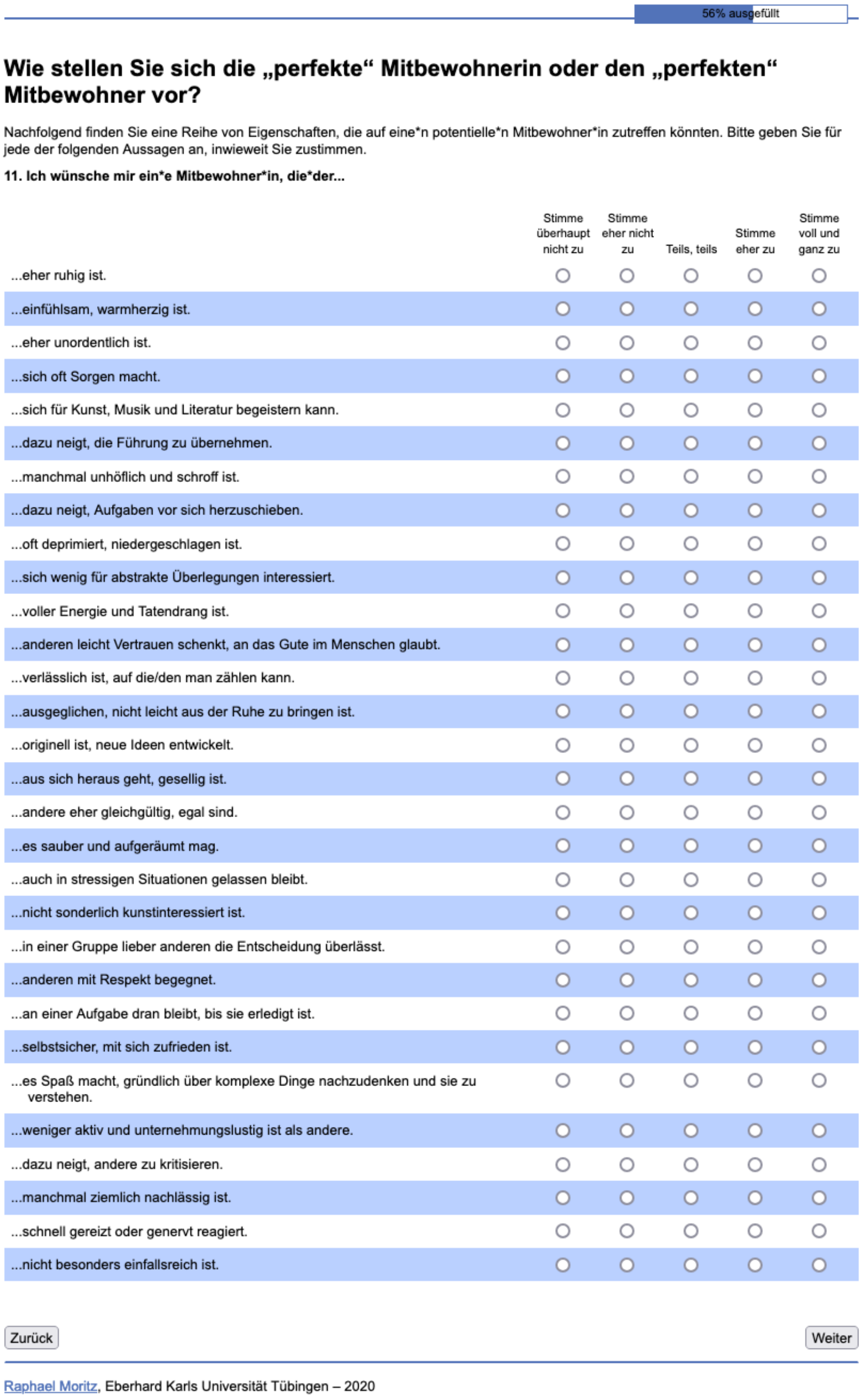
[2] For a comprehensive list of items, see Table D.23.

[3] In this context, a shared apartment is defined as living together with at least one other person who is not related. This definition would also encompass living together with a partner/spouse without being married.

[4] A correlation matrix (see Table D.22) shows that neuroticism/emotional instability is highly (negatively) correlated with conscientiousness as well as agreeableness.

harmoniously with the other roommates was the most crucial factor, followed by the capacity to engage in meaningful conversations, the willingness to participate in shared activities such as cooking, drinking, or watching movies together, and the ability to respect personal property. On average, respondents indicate that working in a similar field, studying a similar subject, having the same level of education, or having a part-time job were not important characteristics to consider when selecting a roommate.

With regard to the preferred gender of potential roommates, 66.7% of respondents indicate that they have no preference. 26.6% of respondents indicate a preference for a female roommate, while 6.7% prefer a male roommate. While respondents indicate that they do not care about the origin or ethnicity of a potential roommate on average, a larger proportion states that a potential roommate should speak the same mother tongue.



56% ausgefüllt

Wie stellen Sie sich die „perfekte" Mitbewohnerin oder den „perfekten" Mitbewohner vor?

Nachfolgend finden Sie eine Reihe von Eigenschaften, die auf eine*n potentielle*n Mitbewohner*in zutreffen könnten. Bitte geben Sie für jede der folgenden Aussagen an, inwieweit Sie zustimmen.

11. Ich wünsche mir ein*e Mitbewohner*in, die*der...

	Stimme überhaupt nicht zu	Stimme eher nicht zu	Teils, teils	Stimme eher zu	Stimme voll und ganz zu
...eher ruhig ist.					
...einfühlsam, warmherzig ist.					
...eher unordentlich ist.					
...sich oft Sorgen macht.					
...sich für Kunst, Musik und Literatur begeistern kann.					
...dazu neigt, die Führung zu übernehmen.					
...manchmal unhöflich und schroff ist.					
...dazu neigt, Aufgaben vor sich herzuschieben.					
...oft deprimiert, niedergeschlagen ist.					
...sich wenig für abstrakte Überlegungen interessiert.					
...voller Energie und Tatendrang ist.					
...anderen leicht Vertrauen schenkt, an das Gute im Menschen glaubt.					
...verlässlich ist, auf die/den man zählen kann.					
...ausgeglichen, nicht leicht aus der Ruhe zu bringen ist.					
...originell ist, neue Ideen entwickelt.					
...aus sich heraus geht, gesellig ist.					
...andere eher gleichgültig, egal sind.					
...es sauber und aufgeräumt mag.					
...auch in stressigen Situationen gelassen bleibt.					
...nicht sonderlich kunstinteressiert ist.					
...in einer Gruppe lieber anderen die Entscheidung überlässt.					
...anderen mit Respekt begegnet.					
...an einer Aufgabe dran bleibt, bis sie erledigt ist.					
...selbstsicher, mit sich zufrieden ist.					
...es Spaß macht, gründlich über komplexe Dinge nachzudenken und sie zu verstehen.					
...weniger aktiv und unternehmungslustig ist als andere.					
...dazu neigt, andere zu kritisieren.					
...manchmal ziemlich nachlässig ist.					
...schnell gereizt oder genervt reagiert.					
...nicht besonders einfallsreich ist.					

Zurück

Weiter

Raphael Moritz, Eberhard Karls Universität Tübingen – 2020

Figure D.20: Screenshot of the Personality Task (First Online Survey)

Table D.21: First Online Survey: Summary Statistics

	(1) Mean	(2) S.D.	(3) Min	(4) Max	(5) Diff. from Midpoint
Extraversion	3.276	0.484	1.500	4.667	0.276*** (0.000)
Agreeableness	4.311	0.400	2.833	5.000	**1.311*** ** (0.000)
Conscientiousness	4.365	0.466	2.333	5.000	**1.365*** ** (0.000)
Neuroticism/Emotional Instability	1.771	0.410	1.000	3.000	**-1.229*** ** (0.000)
Openness	3.679	0.509	1.000	5.000	0.679*** (0.000)
Obs.	768				

Note: The table reports summary statistics for each Big Five personality item on a potential roommates' personality traits. The items are measured on a 5-point Likert scale (1 = do not agree at all, 5 = completely agree). Column 5 shows the results of a one-sample t-test testing the null hypothesis that the mean of the variables is equal to 3 (midpoint of the scale). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D.22: First Online Survey: Correlation Matrix

	(1)	(2)	(3)	(4)	(5)
(1) Extraversion	1				
(2) Agreeableness	0.213***	1			
(3) Conscientiousness	-0.0836*	0.296***	1		
(4) Neuroticism/Emotional Instability	-0.150***	-0.316***	-0.403***	1	
(5) Openness	0.301***	0.220***	0.0591	-0.245***	1

Note: The table reports Pearson correlation coefficients among the Big Five personality traits. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table D.23: First Online Survey: Summary Statistics – Preferences on Characteristics of Potential Roommates

	Mean	S.D.
Does not smoke	3.898	1.371
Experience living in shared apartments	2.982	1.172
Same sex as me	1.962	1.176
As old as me	2.661	1.114
Studies a similar subject	1.715	0.898
Same study level (Bachelor's/Master's)	1.832	1.014
Works in a similar field	1.628	0.814
Has a steady income	2.596	1.228
Has a part-time job	1.847	0.916
Has similar hobbies	2.685	1.041
Has similar interests	3.320	0.987
Has a similar taste in music	2.616	1.089
Fits in with the rest of the roommates	4.615	0.627
Has a well-groomed appearance	3.753	0.963
Goes about her own life	3.776	0.820
Has friends and activities outside the shared apartment	3.987	0.880
With whom I can be friends	3.938	0.891
With whom I can talk	4.316	0.764
With whom I can cook together, have a drink or watch a movie	4.165	0.913
Does not take things out of the fridge that are not her own	4.013	1.098
Obs.	768	

Note: The table reports means and standard deviations for the third section of the survey, asking respondents about the perceived importance of a pre-defined set of characteristics for a potential roommate. The items are measured on a 5-point Likert scale, ranging from 1 = not at all important to 5 = very important.

D.5.2 First Online Pilot Experiment

The findings of the initial online survey (see Section D.5.1) indicate that individuals prefer roommates who are highly conscientious and highly agreeable.[5] In addition, both personality traits are of particular importance within organizational settings (R. Fang et al., 2015; Kluemper & Rosen, 2009). Accordingly, we conduct an online pilot experiment with the objective of gaining insights into the effect of conscientiousness and agreeableness signaling images on social media. The online experiment was conducted with students at the University of Tübingen in July and August 2020 with a total of $n = 467$ respondents.

Survey Design Subjects are randomly assigned to one of the four treatment conditions, which vary in terms of the level (high *vs.* low) of conscientiousness *vs.* agreeableness. Each treatment condition is associated with a social media profile that features images of an individual, designed to manipulate perceptions of the intended personality traits. The images on each profile are compiled by research assistants who are pursuing a degree in psychology with a focus on social media and impression formation.

Subsequently, subjects are instructed to study a screenshot of the given social media profile in great detail and are asked to rate the displayed individual's personality using the Big Five Inventory-2-S, which was also used in the initial survey (see Section D.5.1). Additionally,

[5] Additionally, individuals tend to prefer potential roommates who exhibit low levels of neuroticism/emotional instability as described above. This will be included in a later online pilot study in order to account for a high correlation between agreeableness and neuroticism. Instead of neuroticism, we will then focus on emotional stability in order to align the directions of the effects regarding the high and low levels of the respective traits.

participants are asked to indicate whether, based on the information provided, they would consider inviting the respective individual for a viewing or living with the person. Finally, we asked participants to state their age, gender, and living situation – along with a brief attention check.

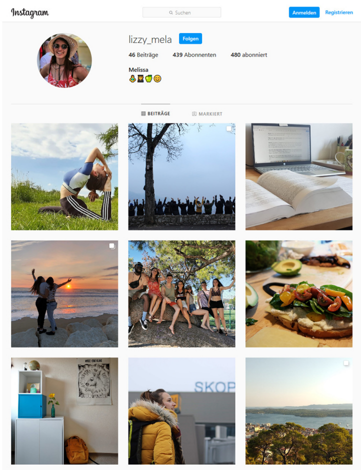
Results 56 observations are excluded from the analysis as they complete the survey too fast or fail to respond correctly to the attention check.[6] The majority of respondents are between the age of 18 and 29, with 70.2% identifying as female. At the time of participation in the survey, 31.3% of respondents are living in a shared apartment, while 23.4% are cohabiting with a partner. None of the subjects indicated to know the displayed person.

The results in Tables D.24 and D.25 indicate that the manipulation of information pertaining to the conscientiousness treatments effectively affected the subjects' perceptions of this specific personality trait exhibited by the displayed individual. With regard to the level of the traits (high *vs.* low), conscientiousness exhibits a large and statistically significant difference between the high and low treatment levels. More importantly, the average marginal effect of the subject's perception of conscientiousness on the conscientiousness treatments is large and statistically significant (see columns 1 and 2 of Table D.25). Furthermore, the magnitude of this effect is the highest of all other personality variables and exhibits the highest level of statistical significance ($p = 0.000$) compared to the other trait variables.

In contrast, the results of the agreeableness treatments do not provide evidence that the relevant personality traits were successfully manipulated. With respect to the level of the traits, the largest and most statistically significant differences are observed for openness and extraversion (see column 6 of Table D.24). Regarding the average marginal effects reported in Table D.25, openness has a larger (negative) and statistically significant ($p = 0.000$) effect on the treatment variable (see columns 3 and 4 of Table D.25) than agreeableness.

With respect to the hypothetical probability of receiving an invitation to a viewing or the possibility of sharing an apartment with the displayed individual, the survey participants assign the highest probability to the high conscientiousness treatment (see column 1 of Table D.26). Furthermore, the difference between high and low levels of conscientiousness is substantial and statistically significant. In line with the aforementioned results, the agreeableness treatment does not demonstrate a statistically significant difference between the high and low levels of the respective treatment. Therefore, we examine adapted versions of social media information and their impact on the perceived personality traits of the individual depicted, with a particular focus on agreeableness, in another online pilot study (see Section D.5.3).

[6] The inclusion of observations that fail to respond correctly to the attention check does not affect the results.



Inwieweit treffen die folgenden Aussagen Deiner Einschätzung nach auf die gezeigte Instagram Nutzerin zu?

Es geht um Deine persönliche Einschätzung aufgrund des Profils. Es gibt keine falschen Antworten.

Die Nutzerin...	stimme gar nicht zu	stimme eher nicht zu	teils, teils	stimme eher zu	stimme voll zu
...ist eher ruhig.	○	○	○	○	○
...ist einfühlsam, warmherzig.	○	○	○	○	○
...ist eher unordentlich.	○	○	○	○	○
...macht sich oft Sorgen.	○	○	○	○	○
...kann sich für Kunst, Musik und Literatur begeistern.	○	○	○	○	○
...neigt dazu, die Führung zu übernehmen.	○	○	○	○	○
...ist manchmal unhöflich und schroff.	○	○	○	○	○

Weiter

Eberhard Karls Universität Tübingen – 2020

33% ausgefüllt

Figure D.21: Screenshot of the Evaluation Task (First Online Pilot Experiment)

Table D.24: First Online Pilot Experiment: Summary Statistics

	High		Low		
	Mean	SD	Mean	SD	Diff.
Conscientiousness Treatment					
Extraversion	3.673	0.503	3.684	0.621	-0.011 (0.886)
Agreeableness	3.847	0.474	3.465	0.553	0.382*** (0.000
Conscientiousness	3.846	0.541	2.630	0.607	1.216*** (0.000)
Neuroticism/Emotional Instability	2.252	0.480	2.724	0.631	-0.472*** (0.000)
Openness	3.325	0.655	2.981	0.678	0.344*** (0.000)
Agreeableness Treatment					
Extraversion	3.332	0.540	3.000	0.668	0.332*** (0.000)
Agreeableness	3.809	0.458	3.636	0.568	0.173** (0.019)
Conscientiousness	3.330	0.565	3.197	0.599	0.133 (0.109)
Neuroticism/Emotional Instability	2.476	0.509	2.732	0.636	-0.257*** (0.002)
Openness	3.321	0.545	3.774	0.539	-0.453*** (0.000)
Obs.	202		209		

Note: The table reports means and standard deviations for each respective personality item, depending on the treatment condition. The items are measured on a 5-point Likert scale (1 = do not agree at all, 5 = completely agree). Column 6 presents the results of a two-sample t-test with equal variances (corresponding p-values in parentheses). * p < 0.1, ** p < 0.05, *** p < 0.01.

Table D.25: First Online Pilot Experiment: Determinants of the Treatment Conditions (Average Marginal Effects)

	(1) High Conscientiousness Treatment	(2) High Conscientiousness Treatment	(3) High Agreeableness Treatment	(4) High Agreeableness Treatment
Extraversion	-0.0400	-0.0381	0.169***	0.163***
	(0.0534)	(0.0516)	(0.0539)	(0.0527)
Agreeableness	-0.150**	-0.152**	0.252***	0.238***
	(0.0651)	(0.0652)	(0.0689)	(0.0689)
Conscientiousness	0.370***	0.374***	0.0787	0.0868
	(0.0277)	(0.0306)	(0.0655)	(0.0635)
Neuroticism/Emotional Instability	-0.116**	-0.123**	0.0203	0.0154
	(0.0495)	(0.0500)	(0.0666)	(0.0657)
Openness	0.0206	0.0202	-0.434***	-0.426***
	(0.0393)	(0.0416)	(0.0406)	(0.0420)
Obs.	211	211	200	200
Respondent Control Variables	No	Yes	No	Yes
Pseudo R^2	0.527	0.535	0.265	0.281

Note: The table reports the results of a probit regression (average marginal effects) of the Big Five personality traits on the high treatment levels for the two treatments as subsamples. Respondent Control Variables include age, female, living situation, and the time the respondent needed to complete the survey. Robust standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table D.26: First Online Pilot Experiment: Hypothetical Callback

	(1) High	(2) Low	(3) Diff.
Conscientiousness	2.292	3.086	0.793***
	(0.956)	(1.110)	(0.000)
Agreeableness	2.563	2.385	-0.178
	(0.916)	(0.978)	(0.187)
Obs.	202	209	

Note: The table reports means and standard deviations of the participant's willingness to live together with the individual, depending on the treatment condition. The item is measured on a 5-point Likert scale (1 = very likely, 5 = very unlikely). Column 3 presents the results of a two-sample t-test with equal variances (corresponding p-values in parentheses). * p < 0.1, ** p < 0.05, *** p < 0.01.

D.5.3 Second Online Pilot Experiment

To create unambiguous and precise treatment conditions that signal high and low levels of conscientiousness and agreeableness, we modify the previously tested social media information (see Section D.5.2), with particular attention to the agreeableness conditions. In collaboration with research assistants and students living in shared apartments, we develop new sets of social media contents. We then conduct a second randomized online pilot experiment to assess whether these revised treatment conditions successfully manipulate participants' beliefs regarding the respective personality traits. We conduct the online experiment with students at the University of Tübingen in July 2023 with a total of $n = 652$ subjects.

Survey Design After a brief introduction and an informed consent page, participants are randomly assigned to one of the four treatment conditions. They receive a screenshot of the entire profile and screenshots of the first ten posts (including descriptions, hashtags, etc.) and are instructed to examine the entire profile and posts in detail. The profiles and posts are identical to those we use in the later field experiments. Participants are then asked to rate the personality traits of the individual displayed on the social media profile using the Big Five Inventory-2-S (see sections D.5.1 and D.5.2). To ensure that participants had not previously taken part in the first online pilot experiment or were not acquainted with the individual depicted (a student at the University of Tübingen), we included a series of control questions.

Subsequently, we ask participants whether they could imagine sharing an apartment with the respective person (hypothetical callback), what motivated this decision (open-ended question), and whether they believe the person regularly fulfills duties within a shared apartment and contributes to its communal life. Finally, participants provided information about their hobbies and demographic characteristics, including gender, age, field of study, living situation, and social media usage.

Results We exclude 64 observations because these participants either report knowing the displayed person, having participated in a similar study, only view the profiles without assessing personality traits, or complete the survey too quickly (Leiner, 2019). Regarding demographic characteristics, the final sample comprise $n = 588$ participants, of whom 74.2% identify as female. The average age is 27.3 years, and participants live with an average of 3.9 other individuals or roommates, with 66.7% residing in a shared apartment. Additionally, 15.9% indicate to be faculty members rather than students. Instagram, the social media platform we use for both experiments, is the most frequently used social media platform among the subjects.

The descriptive results in Table D.27 indicate that the means for conscientiousness and agreeableness are significantly higher in the high conditions compared to the low conditions. Additionally, the magnitude of both personality traits is lowest in the low conditions and highest in the high conditions. Since emotional stability (the opposite of neuroticism) exhibits similar patterns and has a statistically significant large effect on agreeableness (see columns 3 and 4 of Table D.28), we rename the agreeableness conditions to agreeableness/emotional stability, as described above.

Furthermore, the results in Table D.28 indicate that assessing the personality trait of conscientiousness in the high treatment condition yields the largest positive and statistically significant effect among all personality traits (0.409, $p = 0.000$; see column 2 of Table D.28). Additionally, regressing the respective treatment condition on the assessed Big Five personality traits confirms these findings (see column 3 of Table D.29).

For the agreeableness conditions, we observe similar effects. Assessing agreeableness has the largest positive and statistically significant effect on the respective treatment (0.224, $p = 0.000$, see column 4 of Table D.28), as demonstrated in both Table D.28 and Table D.29. However, as mentioned earlier, emotional stability also exerts a large and significant effect on the respective treatment, making it implausible to estimate these effects separately in the main field experiments. Overall, the findings indicate that both conditions – conscientiousness and agreeableness/emotional stability – now successfully manipulate relevant beliefs about the respective personality traits.

Consistent with the initial survey results, the personality traits of agreeableness and conscientiousness (followed by emotional stability) have the strongest effects on hypothetical callbacks (see Table D.30). This suggests that individuals possessing these traits are more likely to be selected as potential roommates (see Section 5.6 on Experiment 2).

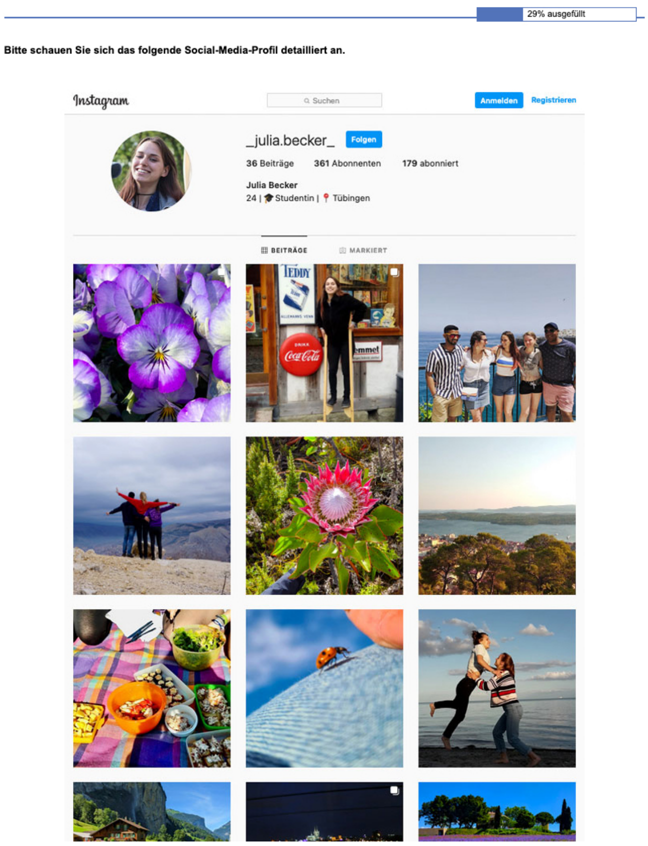


Figure D.22: Screenshot of the Evaluation Task I (Excerpt; Second Online Pilot Experiment)

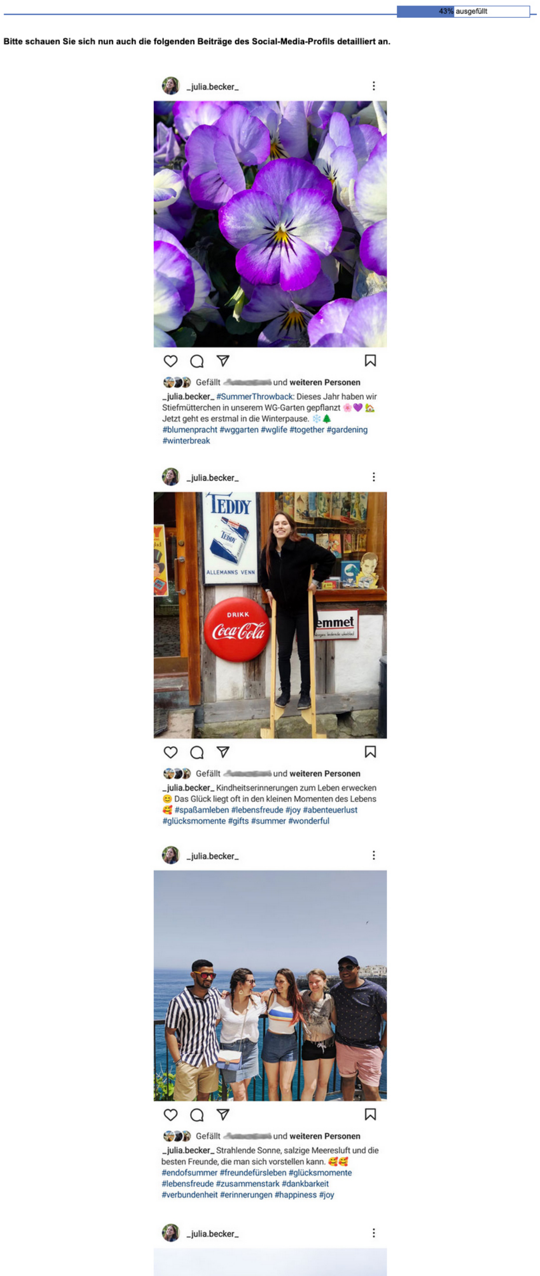


Figure D.23: Screenshot of the Evaluation Task II (Excerpt; Second Online Pilot Experiment)

57% ausgefüllt

1. Wie schätzen Sie die gezeigte Person auf Basis ihres Social-Media-Profils ein?
Uns interessiert Ihr erster Eindruck.

Die gezeigte Person...	Stimme überhaupt nicht zu	Stimme eher nicht zu	Teils, teils	Stimme eher zu	Stimme voll und ganz zu	Ich weiß es nicht
...ist eher ruhig.	○	○	○	○	○	○
...ist einfühlsam, warmherzig.	○	○	○	○	○	○
...ist eher unordentlich.	○	○	○	○	○	○
...macht sich oft Sorgen.	○	○	○	○	○	○
...kann sich für Kunst, Musik und Literatur begeistern.	○	○	○	○	○	○
...neigt dazu, die Führung zu übernehmen.	○	○	○	○	○	○
...ist manchmal unhöflich und schroff.	○	○	○	○	○	○
...neigt dazu, Aufgaben vor sich herzuschieben.	○	○	○	○	○	○
...ist oft deprimiert, niedergeschlagen.	○	○	○	○	○	○
...interessiert sich wenig für abstrakte Überlegungen.	○	○	○	○	○	○
...ist voller Energie und Tatendrang.	○	○	○	○	○	○
...schenkt anderen leicht Vertrauen, glaubt an das Gute im Menschen.	○	○	○	○	○	○
...ist verlässlich, auf sie kann man zählen.	○	○	○	○	○	○
...ist ausgeglichen, nicht leicht aus der Ruhe zu bringen.	○	○	○	○	○	○
...ist originell, entwickelt neue Ideen.	○	○	○	○	○	○
...geht aus sich heraus, ist gesellig.	○	○	○	○	○	○
...betrachtet andere als gleichgültig, egal.	○	○	○	○	○	○
...mag es sauber und aufgeräumt.	○	○	○	○	○	○
...bleibt auch in stressigen Situationen gelassen.	○	○	○	○	○	○
...ist nicht sonderlich kunstinteressiert.	○	○	○	○	○	○
...überlässt in einer Gruppe lieber anderen die Entscheidung.	○	○	○	○	○	○
...begegnet anderen mit Respekt.	○	○	○	○	○	○
...bleibt an einer Aufgabe dran, bis sie erledigt ist.	○	○	○	○	○	○
...ist selbstsicher, mit sich zufrieden.	○	○	○	○	○	○
...hat Spaß dabei, gründlich über komplexe Dinge nachzudenken und sie zu verstehen.	○	○	○	○	○	○
...ist weniger aktiv und unternehmungslustig als andere.	○	○	○	○	○	○
...neigt dazu, andere zu kritisieren.	○	○	○	○	○	○
...ist manchmal ziemlich nachlässig.	○	○	○	○	○	○
...reagiert schnell gereizt oder genervt.	○	○	○	○	○	○
...ist nicht besonders einfallsreich.	○	○	○	○	○	○

2. Kennen Sie die in dem Profil gezeigte Person?

○ Ja
○ Nein

3. Haben Sie bereits an einer Studie teilgenommen, die ein Social-Media-Profil dieser Person zeigte?

○ Ja
○ Nein

Zurück | Weiter

M.Sc. Raphael Moritz, Eberhard Karls Universität Tübingen – 2023

Figure D.24: Screenshot of the Evaluation Task III (Second Online Pilot Experiment)

Table D.27: Second Online Pilot Experiment: Summary Statistics – Randomized Treatments

	High Conscientiousness		Low Conscientiousness		High Agreeableness/ Emotional Stability		Low Agreeableness/ Emotional Stability	
	Mean	Obs.	Mean	Obs.	Mean	Obs.	Mean	Obs.
Extraversion	3.815	101	4.202	104	3.694	103	2.728	100
Agreeableness	**3.874**	103	**3.186**	77	**4.140**	106	**2.062**	123
Conscientiousness	**4.272**	112	**2.146**	89	**3.692**	91	**2.413**	69
Neuroticism	2.212	99	2.660	78	2.186	94	4.180	114
Emotional Stability	**3.788**	99	**3.339**	78	**3.814**	94	**1.820**	114
Openness	3.274	93	2.926	77	3.491	90	3.090	89
Callback	3.169	148	2.273	143	3.245	147	2.020	150
Person Fulfills Duties	4.264	148	2.592	142	3.837	147	2.793	150
Person does not contribute	2.354	147	2.521	142	2.122	147	3.300	150
Female	0.703	148	0.706	143	0.823	147	0.733	150
Male	0.264	148	0.266	143	0.136	147	0.247	150
Diverse	0.0135	148	0.0140	143	0.0272	147	0.0200	150
Age	27.78	145	26.86	136	26.99	141	27.36	141
Household Members	3.764	144	4.263	137	4.284	141	3.224	143
Shared Apartment	0.701	147	0.690	142	0.634	145	0.642	148
Experience Living in Shared App.	4.049	102	4.468	94	3.516	91	4.479	94

Note: The table reports summary statistics for the second online pilot experiment.

Table D.28: Second Online Pilot Experiment: OLS Regression Results

	(1) High Conscientiousness	(2) High Conscientiousness	(3) High Agreeableness/ Emotional Stability	(4) High Agreeableness/ Emotional Stability
Extraversion	-0.100**	-0.0775	0.0753*	0.0495
	(0.0493)	(0.0490)	(0.0416)	(0.0356)
Agreeableness	-0.100**	-0.0887**	0.227***	0.224***
	(0.0403)	(0.0370)	(0.0375)	(0.0405)
Conscientiousness	0.410***	0.409***	0.0177	0.0353
	(0.0192)	(0.0186)	(0.0441)	(0.0412)
Emotional Stability	0.0492	0.0740*	0.161***	0.165***
	(0.0409)	(0.0445)	(0.0435)	(0.0415)
Openness	-0.0423	-0.0577*	-0.0960**	-0.0957**
	(0.0301)	(0.0328)	(0.0390)	(0.0451)
Constant	0.194	0.388	0.343	0.703**
	(0.269)	(0.373)	(0.344)	(0.346)
Obs.	130	121	132	123
R-squared	0.796	0.812	0.809	0.837
Respondent Control Variables	No	Yes	No	Yes

Note: The table reports different OLS regression models with the treatment conditions as dependent variables and the Big Five personality traits as main independent variables. Columns 2 and 4 include respondent control variables. These are: age, female, male, living situation, household members, social media usage, field of studies, and the time the respondent needed to complete the survey. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D.29: Second Online Pilot Experiment: OLS Regression Results II

	(1) Extraversion	(2) Agreeableness	(3) Conscientiousness	(4) Emotional Stability	(5) Openness
Panel A: Conscientiousness					
High Conscientiousness	-0.393***	0.699***	2.135***	0.496***	0.399***
	(0.0680)	(0.0988)	(0.0796)	(0.0923)	(0.113)
Constant	5.031***	3.464***	1.943***	2.322***	2.997***
	(0.246)	(0.690)	(0.384)	(0.616)	(0.533)
Obs.	195	168	189	166	159
R-squared	0.233	0.298	0.801	0.178	0.130
Panel B: Agreeableness/Emotional Stability					
High Agreeableness/ES	0.914***	2.120***	1.305***	1.993***	0.459***
	(0.0966)	(0.0803)	(0.115)	(0.0890)	(0.116)
Constant	2.619***	1.698***	2.201***	4.636***	2.342***
	(0.322)	(0.287)	(0.434)	(0.381)	(0.443)
Obs.	188	212	149	193	165
R-squared	0.391	0.800	0.547	0.773	0.131
Respondent Control Variables	Yes	Yes	Yes	Yes	Yes

Note: The table reports different OLS models regressing the Big Five personality traits on the respective treatment conditions. The abbreviation of "ES" refers to Emotional Stability. For a list of the respondent control variables, see Table D.28. Robust standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table D.30: Second Online Pilot Experiment: OLS Regression Results (Hypothetical Callback)

Hypothetical Callback	(1)	(2)
Extraversion	-0.117*	-0.117*
	(0.0615)	(0.0669)
Agreeableness	0.321***	0.383***
	(0.0828)	(0.0876)
Conscientiousness	0.235***	0.225***
	(0.0576)	(0.0624)
Emotional Stability	0.228***	0.173**
	(0.0739)	(0.0735)
Openness	0.146**	0.157**
	(0.0688)	(0.0767)
Constant	1.428***	0.205
	(0.535)	(0.738)
Obs.	262	244
R-squared	0.560	0.584
Respondent Control Variables	No	Yes

Note: The table reports different OLS regression models with callback as dependent variables and the Big Five personality traits as main independent variable. Column 2 include respondent control variables. For a list of the respondent control variables, see Table D.28. Robust standard errors in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

D.5.4 Second Online Survey

In the second field experiment in the shared housing market (Experiment 2), we include a link to a social media profile within the application text. Although the application text was designed to be as neutral as possible for a student studying in Germany (see Section 5.6), it might still provide information that affect personality assessments. To address this, we introduced an additional baseline treatment that excludes the social media profile from the application for a vacant room (see Section 5.6). Consequently, Experiment 2 comprises five treatment conditions instead of four (high/low conscientiousness and agreeableness/emotional stability) in Experiment 1.

In this third online pilot experiment, we examine which personality traits are associated with the application text that does not include a social media profile (applicable only to Experiment 2). We conduct this last pilot experiment among students on a professional survey platform between July and August 2024 with $n = 341$ subjects.

Survey Design After a brief introduction and an informed consent page, participants are asked to imagine that they live in a shared apartment and are searching for a new roommate online. They are presented with the application text in a standardized format, mirroring the information that advertisers and potential roommates will receive in the subsequent field experiment (see Figure D.25). Participants are instructed to read the text attentively. Subsequently, as in all previous experiments and surveys, participants assess the personality of the fictitious applicant using the Big Five Inventory-2-S (Rammstedt et al., 2020; Soto & John, 2017).

In line with the previous online pilot experiment (see Section D.5.3), we ask participants whether they could imagine sharing an apartment with the applicant (hypothetical callback), what motivated their decision (open-ended question), and whether they believe the person regularly fulfills duties within a shared apartment and contributes to its communal life. We then ask if participants had taken part in a similar study, followed by questions about their hobbies and demographic characteristics, including gender, age, field of study, living situation, and social media usage.

Results We excluded 41 observations from the analysis because these participants completed the survey too fast (Leiner, 2019), did not finish the survey, or had taken part in a similar study. Among the remaining respondents, 80.9% identify as female, with an average age of 23.5 years. On average, participants live with 4.5 other individuals or roommates, and 92.6% reside in a shared apartment for an average of 3.8 years. Only 1% indicated to be faculty members, while the majority are students. Additionally, the vast majority reported using Instagram as their most frequently used social media platform.

The descriptive results in Table D.31 indicate that the mean scores for the Big Five personality traits assessed from the application text predominantly hover around the midpoint of the scale. Specifically, the application text reflects moderate levels of extraversion and openness. Agreeableness and emotional stability exhibit slightly higher mean scores, with deviations of approximately 0.6 and 0.7 points above the midpoint, respectively. However, participants assess the applicant for a vacant room as being above average in conscientiousness, with a score 0.9 points above the midpoint. All differences from the midpoint of the scale are statistically significant. In the analysis of the results of Experiment 2, these values can serve as a benchmark for evaluating the assessed personality of the applicant.

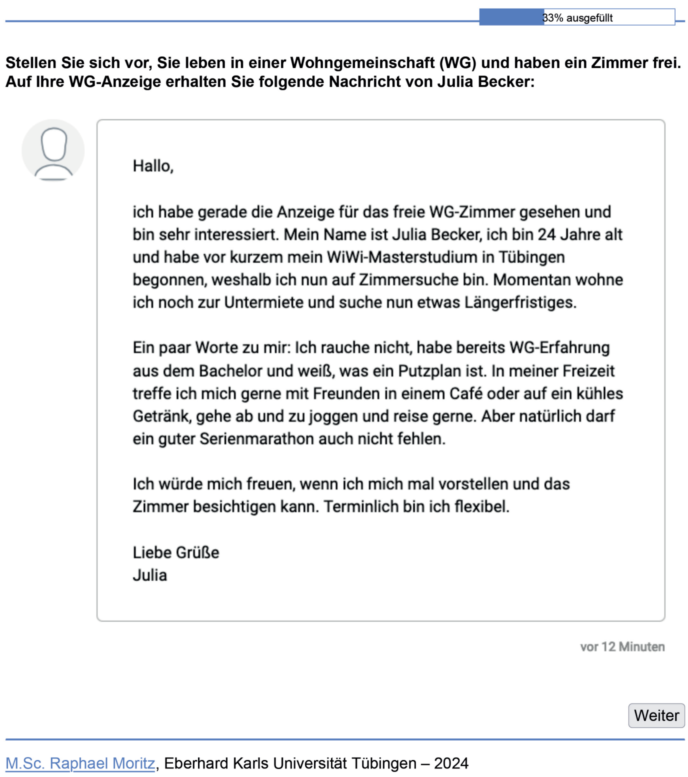


Figure D.25: Screenshot of the Treatment Condition (Second Online Survey)

Table D.31: Second Online Survey: Summary Statistics

	(1) Mean	(2) S.D.	(3) Obs.	(4) Diff. from Midpoint
Extraversion	3.109	0.666	218	0.109** (0.016)
Agreeableness	3.693	0.651	175	0.693*** (0.000)
Conscientiousness	3.949	0.506	191	0.949*** (0.000)
Neuroticism/Emotional Instability	2.362	0.561	149	-0.638*** (0.000)
Emotional Stability	3.638	0.561	149	0.638*** (0.000)
Openness	2.867	0.764	133	-0.133** (0.047)

Note: The table reports summary statistics for each Big Five personality item on a potential roommates' personality traits. The items are measured on a 5-point Likert scale (1 = do not agree at all, 5 = completely agree). Column 4 shows the results of a one-sample t-test testing the null hypothesis that the mean of the variables is equal to 3 (midpoint of the scale). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

D.5.5 Third Online Survey

In order to prevent our social media profiles from being identified as fictitious, inauthentic, or inactive, it is essential to ensure that the profiles exhibit a social network and some level of activity, as evidenced by a sufficient number of friends/subscribers, subscriptions, comments and likes. A paucity of friends/subscribers and likes may be indicative of minimal profile activity or a limited social network, which could suggest a lack of popularity or engagement. Consequently, we ask students from our university to act as friends/subscribers to our fictitious accounts and to like some of the photos with their personal accounts. We conducted this third survey in July and August 2024, right before the start of Experiment 2.

Asking actual humans has important advantages. Primarily, the demographic characteristics of followers align with the narrative of the individual displayed on the profile, making it difficult to detect it as being fictitious. Secondly, an individual who receives a friendship request from one of our profiles (Experiment 1) or the profile link as supplementary information within an application for a vacant room (Experiment 2) may be suspicious if the number of followers is relatively low or if the majority of followers are bot or fake accounts.[7] Thirdly, it is difficult for the social media platform to expose our profiles as fictitious and possibly block them.

Survey Design Firstly, participants are asked to follow a random selection of two out of four profiles. Secondly, we ask respondents to like some of the photos by randomly displaying a set of images of the given profiles with a higher probability for newer photos to appear in order to simulate the effect that newer photos may receive more likes due to a growing social network over time (see Figure D.26). Subjects are instructed to remain subscribed for a period of several months throughout the course of the experiment. In addition, participants are asked to indicate the perceived realism and authenticity of the profiles and to provide any remarks, suggestions, or comments in an open-ended question.

Results A total of $n = 69$ respondents participated in the survey. Regarding the perceived realism and authenticity of the profiles, the majority of respondents (89.4%) rated them as extremely, very, or somewhat realistic. 60.6% rated them as extremely or very realistic. Only a small fraction of 3% indicate that the profiles are not realistic at all (see Table D.32).

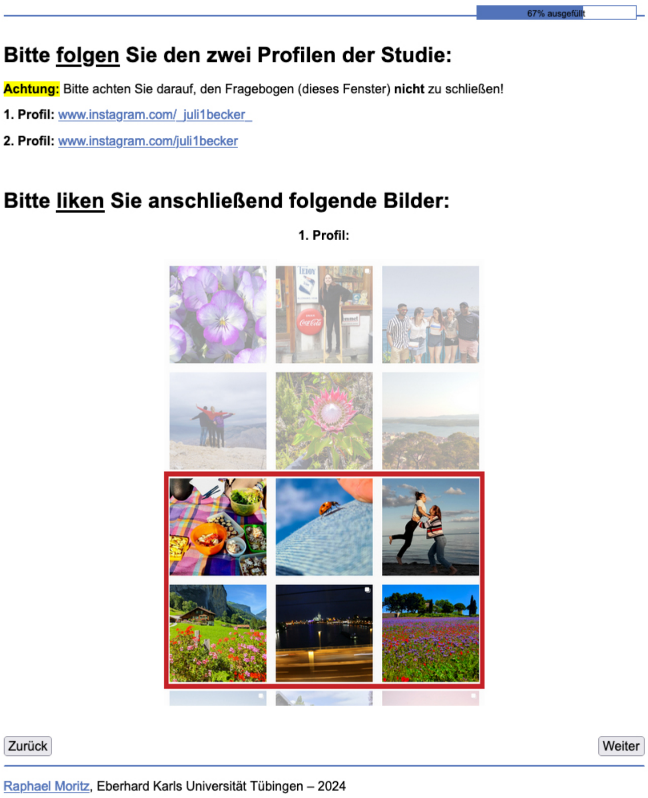


Figure D.26: Screenshot of the Follower Task (Third Online Survey)

[7]While we believe it is unlikely that someone who has received one of our profiles as part of an application or friend request would study the list of friends or subscribers in great detail due to time constraints, we cannot rule out the possibility. Note that in order to access the complete list of friends/subscribers associated with an Instagram profile, the user must first log in. Consequently, this feature is only accessible to registered users (applies to Experiment 2 only).

Table D.32: Third Online Survey: Realism

	All	High Consc.	Low Consc.	High Agreea./ Emot. Stability	Low Agreea./ Emot. Stability
Extremely realistic	16.67 [11]	16.13 [5]	22.58 [7]	14.29 [5]	14.29 [5]
Very realistic	43.94 [29]	48.38 [15]	45.16 [14]	42.89 [15]	40.00 [14]
Somewhat realistic	28.79 [19]	25.81 [8]	22.58 [7]	34.29 [12]	31.43 [11]
Only partly realistic	7.58 [5]	6.45 [2]	9.68 [3]	5.71 [2]	8.57 [3]
Not realistic at all	3.03 [2]	3.23 [1]	0 [0]	2.86 [1]	5.71 [2]

Note: The table presents the proportions of respondents who indicated the respective option if one of the two social media profiles presented belonged to the respective treatment. The numbers in brackets indicate the number of observations, that is, the number of profiles that participants were presented with in the given condition.

D.6 Ethical Considerations

This study explores how different social media profiles, signaling high or low conscientiousness and agreeableness/emotional stability, affect outcomes in two contexts: forming social networks (friend requests) and selecting potential roommates for vacant rooms in shared apartments. By manipulating social media information, we investigate the causal impact of perceived personality traits on selection decisions and social interactions. However, our approach encompasses ethical considerations related to voluntary consent, deception, and the costs imposed on participants, non-participants, and platforms.

As in many field experiments and correspondence studies, participants cannot provide informed consent or choose participation voluntarily. Informing participants would bias their behavior and invalidate results, thus informed consent is not obtained. Moreover, we must use fictitious applicants and social media profiles to ensure genuine responses. Revealing this deception in advance would change subject's behavior and undermine the validity of the findings (Riach & Rich, 2004; Zschirnt, 2019). Although this approach challenges the ideal of voluntary consent, we argue that it is ethically defensible because it is crucial to examine how such online signals affect real-world decisions.

Minimizing Costs to (Non-)Participants and Platforms We aim to minimize costs incurred by individuals exposed to our treatments. The main costs are participants' time in reviewing and responding to a single friend request or room application and possibly inspecting the linked social media profile. Because our study uses a non-matching pairs design, each participant encounters only one request or application, limiting time and costs. We also implement duplicate checks to prevent participants from being treated more than once. In Experiment 2, we reject all responses received within 24 to 48 hours by claiming that the applicant already found a room on short notice. Rejecting responses too quickly might raise suspicion, while waiting too long might impose higher costs.

Additional or psychological costs may arise if participants experience disappointment or frustration after a fictitious applicant withdraws. We acknowledge these concerns but argue that such emotional burdens are minor and transitory, given the context of everyday online interactions. Additionally, we analyze behavior only on an aggregated level, so that individual decision-making and traceability is not possible.

Furthermore, privacy concerns are minimized by exclusively collecting publicly available data that users voluntarily share on their social media profiles or room advertisements. We do not store any information that becomes accessible only, for example, after a friend request is accepted, ensuring that we rely solely on data available to anyone.

Other users of the platforms as well as non-participants in general might be indirectly affected if our fictitious requests/applications displace real ones. But given the high demand for shared rooms (MLP Finanzberatung SE, 2021) and activity on social media platforms, it is unlikely that these effects are substantial. The high volume of friend requests, as well as vacant room ads and applicants per advertisement indicate that our intervention is marginal in comparison to the broader market.

In addition, we introduce only a handful of fictitious profiles on a large-scale social media platform – Instagram has over 1.6 billion users (Datareportal, 2024) – and a major housing website in Germany with millions of monthly unique visitors. Thus, our intervention has negligible effects on server capacity and platform quality. Moreover, we report any spam or bot accounts, as well as fake ads, we encounter to the respective platform which is likely to have a positive effect on platform quality.

Debriefing participants afterwards would disclose the deception and could impose additional time costs and psychological strain, particularly if participants realize their decisions might have been affected by fictitious images. Given these potential burdens and the minimal risks, we do not debrief participants. In our judgment, the disadvantages of added costs, emotional distress, and potentially reduced willingness to participate in future research, outweigh the benefits.

Benefits & Conclusion Our research provides causal evidence on how visual signals of personality traits on social media affect both, the formation of social networks and economic opportunities, such as finding a room. To date, no causal studies exist on how images signaling certain personality traits affect both social acceptance and market outcomes.

Alternative research methods – such as observational and self-report data, laboratory experiments, or large representative samples – would not reliably identify causal effects of perceived personality traits on real decision-making. Observational data cannot isolate causality, and laboratory environments risk biases like the Hawthorne effect. Field experiments using deception are necessary for generating externally valid, causal insights on the role of social media signals in important real-world decisions.

Our results not only inform academic debates but also contribute to public discourse on the significance of online information. Understanding these dynamics is increasingly important as social media shapes political participation, well-being, mental health, and information dissemination (Acquisti et al., 2015; Allcott et al., 2020; Allcott & Gentzkow, 2017).

Although our research design imposes certain costs to participants and raises important ethical concerns, we argue that the societal value and significance of our research questions justify these trade-offs. Moreover, as outlined above, we have implemented several measures to minimize costs for participants, non-participants, and platforms. Ultimately, the benefits of obtaining robust causal evidence on the effect of social media information outweigh the costs imposed by our field experiments.

Appendix E

Appendix Chapter 6

E.1 Variable Definitions

Compensation variables:

- *Base* $salary_{i,t}$: Includes all contractually agreed on fixed compensation components, paid regardless of business, company-wide and/or individual performance. (source: hkp, own data collection)
- $RLTI_{i,t}$: Realized long-term incentive compensation defined as the sum of all variable compensation elements (in Euros) based on a performance period of more than one year. The following plan types are considered: cash and equity deferrals, share matching plans, multi-year bonuses, restricted stock, performance shares, and stock options. In case a long-term incentive plan (LTI) is not granted annually, but upfront for a multi-year period, the grant value is distributed over this period and the pro-rated value is shown for each year. (source: hkp, own data collection)
- $RSTI_{i,t}$: Realized short-term incentive compensation defined as the sum of all variable compensation elements (in Euros) with a performance period of one year paid to executive i in fiscal year t. Payment is made in cash at or shortly after the end of the performance period. (source: hkp, own data collection)
- $TSTI_{i,t}$: Target short-term incentive compensation defined as the amount (in Euros) that the firm targets to pay to executive i in fiscal year t if the executive fully meets her targets, i.e., neither over- nor underperforms. In practice, firms typically report $TSTI_{i,t}$ as a proportion of base salary or as a proportion of a bonus cap/STI cap. Occasionally, a target range is reported instead, in which case we take the midpoint. (source: hkp, own data collection)
- $w_{i,t}^{B,ESG}$: Total weight of all *binding* ESG performance metrics of executive i in fiscal year t. In the case of binding metrics, the firm commits to consider target achievement with a weight that is communicated at the beginning of fiscal year t. (source: own data collection)
- $w_{i,t}^{B,nESG}$: Total weight of all *binding* non-ESG performance metrics of executive i in fiscal year t. In the case of binding metrics, the firm commits to consider target achievement with a weight that is communicated at the beginning of fiscal year t. (source: own data collection)
- $w_{i,t}^{D}$: Total weight of all *discretionary* performance metrics of executive i in fiscal year t, calculated as $w_{i,t}^{D} = 1 - w_{i,t}^{B,ESG} - w_{i,t}^{B,nESG}$. Ex ante, executive i knows that the firm will

consider the achievement of all discretionary metrics jointly with weight $w_{i,t}^{D}$ in the calculation of compensation at year-end. Specific weights for individual discretionary metrics are not known. (source: own data collection)

- $f_{i,t}$: Overall target fulfillment rate for all performance metrics together, defined as the weighted average of specific target fulfillment rates for individual binding metrics and all discretionary metrics combined (Equation 6.2). A value of $f_{i,t}$ =100% means that executive i has fully met her targets, i.e., neither over- nor underperformed in fiscal year t. Values reported in the tables are values as reported in firms' annual statements, i.e., after discretionary adjustments by the supervisory board or compensation committee. (source: own data collection)
- *STI hurdle* $Y/N_{i,t}$: Dummy variable indicating whether a firm reports using a lower threshold for realized STI. If actual performance falls below this lower threshold, realized $RSTI_{i,t}$ is set to a minimum value, typically zero. (source: own data collection)
- *LTI hurdle* $Y/N_{i,t}$: See *STI hurdle* $Y/N_{i,t}$, but for LTI. (source: own data collection)
- *STI cap* $Y/N_{i,t}$: Dummy variable indicating whether STI amount is capped at an upper threshold. (source: own data collection)
- *LTI cap* $Y/N_{i,t}$: See *STI cap* $Y/N_{i,t}$ but for LTI. (source: own data collection)
- *STI board discretion allowed* $Y/N_{i,t}$: Dummy variable indicating whether the supervisory board and/or compensation committee retains the right to subjectively adjust the STI amount ex post (e.g., to account for unpredicted exogenous events). (source: own data collection)
- *LTI board discretion allowed* $Y/N_{i,t}$: See *STI board discretion allowed* $Y/N_{i,t}$, but for LTI. (source: own data collection)
- *STI ex post multiplier (ESG/nESG)* $Y/N_{i,t}$: Dummy variable indicating whether STI has been adjusted ex post according to a multiplier determined by the supervisory board / compensation committee. (source: own data collection)
- *LTI ex post multiplier (ESG/nESG)* $Y/N_{i,t}$: See *STI ex post multiplier (ESG/nESG)* $Y/N_{i,t}$, but for LTI. (source: own data collection)
- *STI partly deferred* $Y/N_{i,t}$: Dummy variable indicating whether parts of the STI amount were deferred, i.e., paid out in later periods. (source: own data collection)
- *STI deferral period*$_{i,t}$: In case of deferral, this variable denotes the period in years over which the STI is to be deferred. (source: own data collection)
- *Board exercised discretion* $Y/N_{i,t}$: Dummy variable indicating whether supervisory board and/or compensation committee has ultimately exercised its right to subjectively adjust the STI amount ex post. (source: own data collection)

Executive characteristics:

- *Age*$_{i,t}$: Age of the executive in years. (source: hkp, own data collection)
- *Tenure*$_{i,t}$: Tenure of the executive in years. (source: hkp, own data collection)
- *Female* $Y/N_{i,t}$: Gender of the executive. (source: hkp, own data collection)
- C-suite position: (source: own data collection)

- *CEO*: Chief Executive Officer
- *CFO*: Chief Financial Officer
- *CHRO*: Chief Human Resources Officer
- *COO*: Chief Operating Officer
- *Other specialists*: All other expert C-suite positions such as chief marketing officer, chief sales officer, chief legal officer, etc.
- *Divisional/regional head*: Executives who are chairing a division, geographic region, or are responsible for a product segment

Company-level variables (all accessed via LSEG Data & Analytics):

- *Market capitalization*$_{f,t}$: Market price year-end multiplied with common shares outstanding (WC08001) (source: Worldscope)
- *Total assets*$_{f,t}$: Total assets (WC02999) (source: Worldscope)
- *EBITDA*$_{f,t}$: Earnings before interest, taxes, depreciation, and amortization represented by sum of EBIT and total depreciation and amortization value for the period (TR.EBITDAActValue) (source: Refinitiv Eikon)
- *EBIT*$_{f,t}$: Earnings before interest & taxes represented by difference between total revenues and total operating expenses (TR.EBITActValue) (source: Refinitiv Eikon)
- *Net income*$_{f,t}$: Income after all operating and non-operating income and expense, reserves, income taxes, minority interest, and extraordinary items (TR.InvtrNetIncome) (source: Refinitiv Eikon)
- *Book-to-market ratio*$_{f,t}$: Book value of common equity (WC03501) divided by market value of equity (WC08001) (source: Worldscope)
- *Book leverage*$_{f,t}$: Total debt (sum of long- and short-term debt) (WC03255) divided by total assets (WC02999) (source: Worldscope)
- *Net PPE/total assets (tangibility)*$_{f,t}$: Gross property, plant, and equipment (PPE) less accumulated reserves for depreciation, depletion, and amortization (WC02501) divided by total assets (WC02999) (source: Worldscope)
- *Return on assets*$_{f,t}$: ((Net income – bottom line + (interest expense on debt-interest capitalized) * (1 - tax rate))) / Average of last year's and current year's total assets (WC08326) (source: Worldscope)
- *Return on equity*$_{f,t}$: (Net income - bottom line - preferred dividend requirement) / Average of last year's and current year's common equity (WC08301) (source: Worldscope)
- *Dividends/earnings*$_{f,t}$: Dividends per share (WC05101) divided by earnings per share (WC05201) (source: Worldscope)
- *Stock (total investment) return*$_{f,t}$: Stock return of the company divided by last year's market value (WC08801) (in percent, source: Worldscope)
- *Historical stock-to-accounting volatility*$_{f}$: Standard deviation of total investment return (WC08801) divided by standard deviation of return on assets (WC08326) between the years 2008 and 2012 (source: Worldscope)

- *Board independence*$_{f,t}$: Share of independent board members (TR.AnalyticIndepBoard) (source: Refinitiv Eikon)
- *Female board membership*$_{f,t}$: Share of female board members (TR.AnalyticBoardFemale) (source: Refinitiv Eikon)
- *Institutional ownership*$_{f,t}$: Share of company's equity owned by institutional investors. Institutional investors include the following categories of financial institutions: investment advisors, hedge funds, pension funds, banks and trusts, insurance companies, sovereign wealth funds, venture capital, and private equity (Shareholders History Report) (source: Refinitiv Eikon)
- *Block ownership*$_{f,t}$: Share of owners that own more than 10% of company's equity (Shareholders History Report) (source: Refinitiv Eikon)
- *Emission pledge Y/N*$_{f,t}$: Company policy to improve emission reduction (TR.PolicyEmissions) (source: Refinitiv Eikon)
- CO_2 *(Scope 1)*$_{f,t}$: Direct of CO2 and CO2 equivalents emissions in tonnes from sources owned or controlled by the company (scope 1 emissions). The following gases are relevant: carbon dioxide, methane, nitrous oxide, hydrofluorocarbons, perfluorinated compound, sulfur hexafluoride, nitrogen trifluoride. Emission classifications based on greenhouse gas protocol (GHG) (ENERDP024) (source: Datastream)
- *Historical Log(Avg.* CO_2*)*$_f$: Natural log of average CO_2 *(Scope 1)*$_{f,t}$ between 2008 and 2012 (source: Datastream)
- *Industry*$_f$: Industry classification is based on the primary Global Industry Classification Standard (GICS) sector code, which we aggregate to six major industry sectors comprising Consumers, Energy & Utilities, Financials, Industrials & Materials, Health, and Information & Communication Technology (ICT).

E.2 Overview – Firms in Sample

Table E.1: List of Sample Firms

Firm	Country	Sample period (FY)
ABB	CH	2013-2020
Adidas	DE	2013-2020
Ahold Delhaize	NL	2016-2020
Air Liquide	FR	2013-2020
Airbus	NL	2013-2020
Allianz	DE	2013-2020
Amadeus IT	ES	2013-2020
Anheuser-Busch InBev	BE	2013-2020
ASML	NL	2013-2020
Assicurazioni Generali	IT	2013-2020
AstraZeneca	GB	2013-2020
AXA	FR	2013-2020
Banco Santander	ES	2013-2020
Barclays	GB	2013-2020
BASF	DE	2013-2020
Bayer	DE	2013-2020
BG Group	GB	2013-2015
BHP Group	AU/GB	2013-2020
BMW	DE	2013-2020
BNP Paribas	FR	2013-2020
British American Tobacco	GB	2013-2020
Carrefour	FR	2013-2020
Credit Suisse	CH	2013-2020
CRH	IE	2013-2020
Daimler	DE	2013-2020
Danone	FR	2013-2020
Deutsche Bank	DE	2013-2020
Deutsche Börse	DE	2013-2020
Deutsche Post	DE	2013-2020
Deutsche Telekom	DE	2013-2020
Diageo	GB	2013-2020
E.ON	DE	2013-2020
Enel	IT	2013-2020
ENGIE	FR	2013-2020
Eni	IT	2013-2020
EssilorLuxottica	FR	2013-2020
Flutter Entertainment	IE	2016-2020
Fresenius	DE	2013-2020
GlaxoSmithKline	GB	2013-2020
Glencore	JE	2013-2020
HSBC	GB	2013-2020
Iberdrola	ES	2013-2020
Imperial Brands	GB	2013-2020
Inditex	ES	2014-2020
ING Groep	NL	2013-2020
Intesa Sanpaolo	IT	2013-2020
Kering	FR	2013-2020
KONE	FI	2013-2020
Linde SE	IE	2018-2020
Lloyds Banking	GB	2013-2020
L'Oréal	FR	2013-2020
LVMH	FR	2013-2020
Münchener Rück	DE	2013-2020
National Grid	GB	2013-2020
Nestlé	CH	2013-2020

Continued on next page

Table E.1 – continued from previous page

Firm	Country	Sample period (FY)
Nokia	FI	2013-2020
Novartis	CH	2013-2020
Novo Nordisk	DK	2013-2020
Orange	FR	2013-2020
Pernod Ricard	FR	2013-2020
Philips	NL	2013-2020
Prosus	NL	2020
Prudential	GB	2013-2020
Reckitt Benckiser	GB	2013-2020
RELX Group	GB	2013-2020
Repsol	ES	2013-2020
Richemont	CH	2013-2020
Rio Tinto	AU/GB	2013-2020
Roche	CH	2013-2020
Royal Dutch Shell	GB	2013-2020
RWE	DE	2013-2020
Safran	FR	2013-2020
Saint-Gobain	FR	2013-2020
Sanofi	FR	2013-2020
SAP	DE	2013-2020
Schneider Electric	FR	2013-2020
Siemens	DE	2013-2020
Société Générale	FR	2013-2020
Standard Chartered	GB	2013-2020
Syngenta	CH	2013-2016
Telefónica	ES	2013-2020
Total	FR	2013-2020
UBS	CH	2013-2020
Unibail-Rodamco-Westfield	FR	2018-2020
UniCredit	IT	2013-2020
Unilever Group	GB	2013-2020
Vinci	FR	2013-2020
Vivendi	FR	2013-2020
Vodafone	GB	2013-2020
Volkswagen	DE	2013-2020
Vonovia	DE	2015-2020
Zurich Insurance	CH	2013-2020

Notes: Our sample includes all firms that have been listed for at least ten days on either the EUROSTOXX 50 or the STOXX Europe 50 (or both), between December 31, 2014 and December 31, 2020. Index composition includes changes up until FY2020.

E.3 ESG Classification According to LSEG

The following table shows the underlying taxonomy which we use to classify the non-financial metrics in executive compensation contracts with respect to the E, S, and G dimension and its subcategories.

Table E.2: ESG Classification

Category	**Subcategory**	**Exemplary metrics**
Environmental	Emission	Emissions (scope 1-3), GHG protocol
		Emission reduction target
		Internal carbon pricing
		Waste & waste recycling
		Biodiversity
		Environmental management systems
	Innovation	Product innovation (environmental, organic, eco-design products)
		Green revenues (from environmental products)
		Environmental R&D expenditures
		Environmental assets under management
		Environmental project financing
	Resource use	Water
		Energy
		Toxic chemicals
		Renewable energy
		Sustainable packing policy
		Environmental supply chain management
Social	Community	Fair competition
		Bribery and corruption policy
		Whistleblower protection
		Donations
		Product Sales at Discount to Emerging Markets
		Corporate Responsibility Awards
		Crisis Management Systems
	Human rights	Human rights policy
		Child labor policy
		Ethical Trading Initiative
		Freedom of Association policy
	Product responsibility	Responsible marketing
		Customer Health & Safety policy
		Product quality
		Quality management system
		Customer satisfaction
		Healthy foods or products
		Retailing responsibility
		Revenues from alcohol, tobacco, gambling, armaments

Continued on next page

Table E.2 – continued from previous page

Category	Subcategory	Exemplary metrics
Social	Workforce	Diversity, equity, and inclusion
		Career development and training
		Employee satisfaction and turnover
		Employees with disabilities
		Female representation (top) management
		Workplace accidents
		Working conditions
		Gender pay gap
		Employee health and safety policy
Governance	CSR strategy	CSR strategy
		Stakeholder engagement
		ESG and GRI reporting scope
		UNPRI Signatory
	Management	Board independence and diversity
		Committees (Audit, compensation, nomination)
		Board size
		Succession plans
		Executive Compensation Policy
	Shareholders	Shareholder rights policy
		Shareholders Vote on Executive Pay
		Shareholder Approval Significant Transactions
		Takeover defences

E.4 Exemplary Disclosure of ESG-Related Metrics from Company Reports

This section highlights examples of the disclosure of ESG-related metrics in executive STI compensation contracts, differentiating between low-disclosure and high-disclosure cases w.r.t. to the richness and salience of information on ESG metrics.

Annual variable remuneration	Cap 100% of the fixed remuneration	The annual variable remuneration is designed to align the executive officer's remuneration with the Group's annual performance and to promote the implementation of its strategy year after year. The Board of Directors strives to encourage the executive officer both to maximise performance for each financial year and to ensure that it is repeated and regular year-on-year. Annual variable remuneration can amount to a maximum of 100% of the fixed remuneration.

Criteria for assessment of performance for 2020	Weightings
• **Financial criteria**	**60%**
• Growth in comparable sales as compared to the budget	*15%*
• Growth in market share as compared to the main competitors	*15%*
• Growth in operating profit as compared to 2019	*10%*
• Growth in net earnings per share as compared to 2019	*10%*
• Growth in cash flow as compared to 2019	*10%*
• **Extra-financial and qualitative criteria**	**40%**
• *Quantifiable criteria: (% allocated equally among the following criteria)* • CSR (*Sharing Beauty With All* programme) • Human Resources: gender parity, development of talented employees, access to training • Digital Development	*25%*
• *Individual qualitative performance:* Management, image, company reputation, dialogue with stakeholders.	*15%*

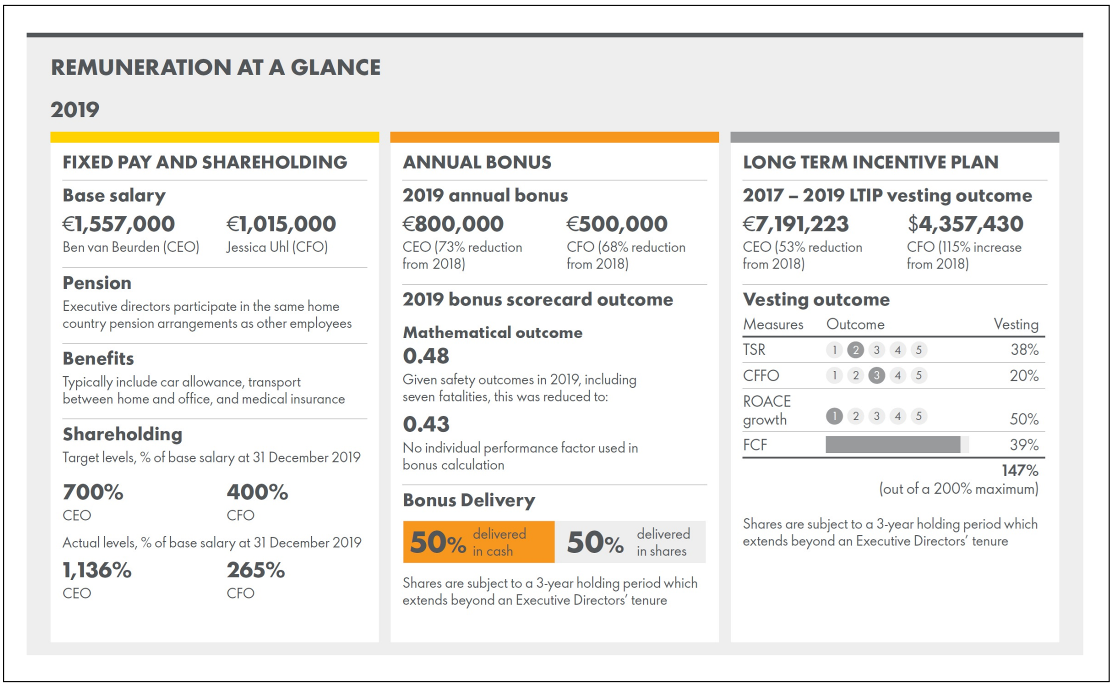
The quantifiable, financial (60%) and non-financial (25%) criteria account for 85% of annual variable remuneration. The weighting of each of these criteria, both financial, extra-financial and qualitative, and the targets to be met were set at the start of the year and communicated to the executive officer. The assessment is made without offsetting among criteria. Pursuant to Article L. 225-100 of the French Commercial Code, payment of the annual variable remuneration is conditional on approval by the Annual General Meeting called to approve the 2020 financial statements.

Figure E.1: L'Oréal 2019 Annual Financial Report (prototype of an STI contract with numerous ESG-related metrics, which are simultaneously also binding)

2020 NON-FINANCIAL AND QUALITATIVE TARGETS
TABLE OF NON-FINANCIAL AND QUALITATIVE ACHIEVEMENTS

CSR criteria: Sharing Beauty with All programme	2020 results	2019 results
The Sharing Beauty with All programme was launched in October 2013 by Mr Jean-Paul Agon. It structures the Group's CSR strategy and sets ambitious targets to be met by 2020. This project consists of four pillars, for which the 2020 achievements are set out in detail in Chapter 4 of this document.		
"Innovating Sustainably"		
• Improved environmental or social impact for 100% of our products.	• 96% of new products that have been screened have an improved environmental or social profile.	85%
"Producing Sustainably" (plants and distribution centres)		
• -60% on CO_2 emissions.	• -81% on CO_2 emissions (absolute value).	-78%
• -60% in water consumption.	• -33% water consumption (per finished product unit).	-51%
• -60% waste generation.	• -37% waste production (by finished product unit) The objective of 0% waste from plants and distribution centres sent to landfill was met from 2018.	-35%
"Living Sustainably"		
• Each brand will have assessed its environmental and social footprint.	• 89% of brands have assessed their impact.	89%
• Each brand will have reported on its progress and associated consumers with its commitments.	• 79% of the brands conducted a consumer awareness initiative.	57%
"Developing Sustainably"		
• With the employees (*L'Oréal Share & Care* programme): 100% of L'Oréal employees around the world will have access to healthcare coverage and social protection in 2020.	• 96% of the Group's permanent employees have access to healthcare coverage reflecting the best practices in their country of residence.	94%
	92% of the Group's permanent employees benefit from financial protection in the event of life-changing accidents, such as death or permanent disability.	91%
• With strategic suppliers.	• 99% of the Group's strategic suppliers carried out a self-assessment of their Sustainable Development policy.	96%
• With communities.	• Access to work for 100,905 people.	90,635
Human Resources criteria	**2020 results**	**2019 results**
Gender Balance		
• Improving gender balance, in particular at the level of senior management positions.	• 26% of Executive Committee members are women	30%
	• 49% of key positions held by women	47%
	• 2020 Equileap ranking:	
	France: No. 1	No. 2 in Europe
	International: No. 4	
	• For the fourth consecutive year, L'Oréal is ranked in the Top 100 among the 325 companies of the Bloomberg Gender-Equality Index 2021.	Top 100
Talent Development		
• Positive policy results for the recruitment of both experienced and more junior talented employees, and talent development all over the world, in order to favour the emergence of local talent.	• 10th place in the Universum global ranking (business schools).	12th place
	L'Oréal is the only French and European company in the Top 10.	
• Attractive, targeted, digital employer communication.	• Strong presence on social networks: 3.2 million followers on LinkedIn.	2.3 million
	• 3rd place in the PotentialPark global ranking (recruiting/social media).	No. 2
Access to training		
• 100% of employees will receive training once a year starting in 2020.	• 100% of employees received training in 2020.	96%
	Over 580,000 hours of digital training.	202,000 hours

Figure E.2: L'Oréal 2020 Annual Financial Report (prototype of an STI contract with numerous ESG-related metrics, which are simultaneously also binding)



REMUNERATION AT A GLANCE

2019

FIXED PAY AND SHAREHOLDING

Base salary

€1,557,000 Ben van Beurden (CEO)

€1,015,000 Jessica Uhl (CFO)

Pension

Executive directors participate in the same home country pension arrangements as other employees

Benefits

Typically include car allowance, transport between home and office, and medical insurance

Shareholding

Target levels, % of base salary at 31 December 2019

700% CEO

400% CFO

Actual levels, % of base salary at 31 December 2019

1,136% CEO

265% CFO

ANNUAL BONUS

2019 annual bonus

€800,000 CEO (73% reduction from 2018)

€500,000 CFO (68% reduction from 2018)

2019 bonus scorecard outcome

Mathematical outcome

0.48

Given safety outcomes in 2019, including seven fatalities, this was reduced to:

0.43

No individual performance factor used in bonus calculation

Bonus Delivery

50% delivered in cash

50% delivered in shares

Shares are subject to a 3-year holding period which extends beyond an Executive Directors' tenure

LONG TERM INCENTIVE PLAN

2017 – 2019 LTIP vesting outcome

€7,191,223 CEO (53% reduction from 2018)

$4,357,430 CFO (115% increase from 2018)

Vesting outcome

Measures	Outcome	Vesting
TSR	1 **2** 3 4 5	38%
CFFO	1 2 **3** 4 5	20%
ROACE growth	**1** 2 3 4 5	50%
FCF		39%
		147% (out of a 200% maximum)

Shares are subject to a 3-year holding period which extends beyond an Executive Directors' tenure

Figure E.3: Royal Dutch Shell 2019 Annual Financial Report (prototype of an STI contract with ESG metrics which are all binding, no subjective component

Notes to the single total figure of remuneration for executive directors table (audited)

Annual bonus

The Annual bonus operated in line with the policy as disclosed on page 156.

Determination of the 2019 annual bonus

The table below summarises the 2019 annual bonus scorecard measures including their weightings, targets and outcomes. The mathematical scorecard outcome for 2019 was 0.48. Please refer to pages 135-136 for a commentary on the scorecard outcome.

After reviewing the mathematical scorecard outcome, and considering the context of wider company performance for the year, the REMCO exercised discretion to adjust the scorecard result downwards to 0.43. This downwards adjustment was to reflect the seven fatalities under Shell operational control during the year.

Accordingly, the REMCO determined a final bonus outcome of €800,000 for the CEO which is 41% of target and 21% of maximum. This is a 73% reduction from 2018. The REMCO determined a final bonus outcome of €500,000 for the CFO which is 41% of target and 21% of maximum. This is a 68% reduction from 2018.

2019 annual bonus outcome (audited) [A][B]

Measures	Weight (% of scorecard)	Threshold	Target set	Outstanding	Result achieved	Score (0-2)
Cash flow from operating activities ($ billion)	**30%**	44	50	56	42	**0**
Operational excellence	**50%**					**0.72**
Production (kboe/d)	12.5%	3,647	3,760	3,873	3,665	0.16
LNG liquefaction volumes (mtpa)	12.5%	35.3	36.4	37.4	35.6	0.23
Refinery and chemical plant availability (%)	12.5%	88.4	90.4	92.4	90.8	1.20
Project delivery on schedule (%)	6.25%	60	80	100	90	1.50
Project delivery on budget (%)	6.25%	105	100	95	99	1.10
Sustainable development	**20%**					**0.59**
Total recordable case frequency (injuries/million hours)	5%	0.9	0.7	0.5	0.9	-
Operational Tier 1 and 2 process safety events (number)	5%	145	115	85	130	0.50
Upstream and Integrated Gas GHG intensity (tonnes of CO_2 equivalent/tonne of hydrocarbon production available for sale)	4%	0.176	0.168	0.160	0.168	1.00
Refining GHG intensity (tonnes CO_2 equivalent per Solomon's Utilized Equivalent Distillation Capacity (UEDC™))	4%	1.11	1.06	1.01	1.06	1.00
Chemicals GHG intensity (tonnes CO_2 equivalent/tonne of petrochemicals production)	2%	1.10	1.00	0.90	1.04	0.60
	100%					
Mathematical scorecard outcome						**0.48**
Adjusted scorecard outcome						**0.43**

[A] These metrics measure the effectiveness with which we operate our assets and portfolio base, assessed against our operational business plan. Shell's longer-term strategic ambitions are measured in the LTIP metrics.
[B] Scorecard targets are based on Shell's annual operating plan and increase or decrease year-on-year. In 2019, target refinery and chemical plant availability was lower and target GHG emission intensities higher than 2018, due to planned business activities, reflecting scheduled maintenance and expected market conditions, and portfolio developments.
[C] In external disclosure, we may use an alternative performance measure, i.e. CFFO excluding Working Capital, to describe the cash flow generation from our operations without the effect of working capital changes.

2019 bonus outcome calculation

Ben van Beurden

Target bonus:

€1,557,000 (base salary) x 125% = **€1,946,250**

×

2019 scorecard result 0.43

=

€800,000 [A]
(51% of base salary)
↓ 73% reduction from 2018

Figure E.4: Royal Dutch Shell 2019 Annual Financial Report (prototype of an STI contract with ESG metrics which are all binding, no subjective component

Variable compensation of the Senior Executive Committee

Executives also benefit from a short-term cash incentive plan and awards granted under the Group's share option and PSU plans. The Committee considers these components in total to ensure there is an appropriate balance between reward for short-term success and long-term retention. Targets used to determine the payout levels for both the variable short-term incentives and the variable long-term incentives are considered by the Committee on an annual basis prior to the start of the next financial year. The Group does not provide for any transaction-specific success fees for its executives.

Short-term cash incentives

Short-term incentives are paid in cash annually and relate to performance in the previous financial year.

The determination of the level of short-term cash incentive comprises both quantitative and qualitative components, each with a pre-set target and a maximum percentage of base salary. The mix of quantitative and qualitative targets are aligned with the Group's business priorities for the year ahead, encouraging individual creativity and delivering continued profit growth and value creation. The short-term incentive target is set at 75% of base salary, with a maximum cap of 150% of base salary.

The quantitative component of the short-term cash incentive is assessed on actual Group or Maison turnover, operating profit and cash generation, being operating cash flow after capital expenditure and lease payments, compared against the current year's budget. Each of these three measures has equal weighting in the calculation. The impact of the Covid-19 outbreak on results for the year has therefore reduced the percentage achievement of quantitative objectives. The qualitative component is assessed on performance against both individual and collective strategic targets, measuring the contribution to creativity, team-building and succession-planning, among other elements.

The total incentive awards achieved represented on average 56% of base salary. The individual figures for the Group's executive directors are as follows:

	Quantitative (% of salary)		Qualitative (% of salary)		Total (% of salary)	
	Target	Achieved	Target	Achieved	Target	Achieved
Nicolas Bos	*41%*	37%	*34%*	38%	*75%*	75%
Burkhart Grund	*41%*	12%	*34%*	36%	*75%*	48%
Sophie Guieysse	*41%*	12%	*34%*	35%	*75%*	47%
Jérôme Lambert	*41%*	12%	*34%*	34%	*75%*	46%
Cyrille Vigneron	*41%*	29%	*34%*	32%	*75%*	61%

Figure E.5: Richemont 2020 Annual Financial Report (prototype of an STI contract with only a few ESG metrics which are also only discretionary, i.e., non-binding metrics)

E.5 Tables & Figures

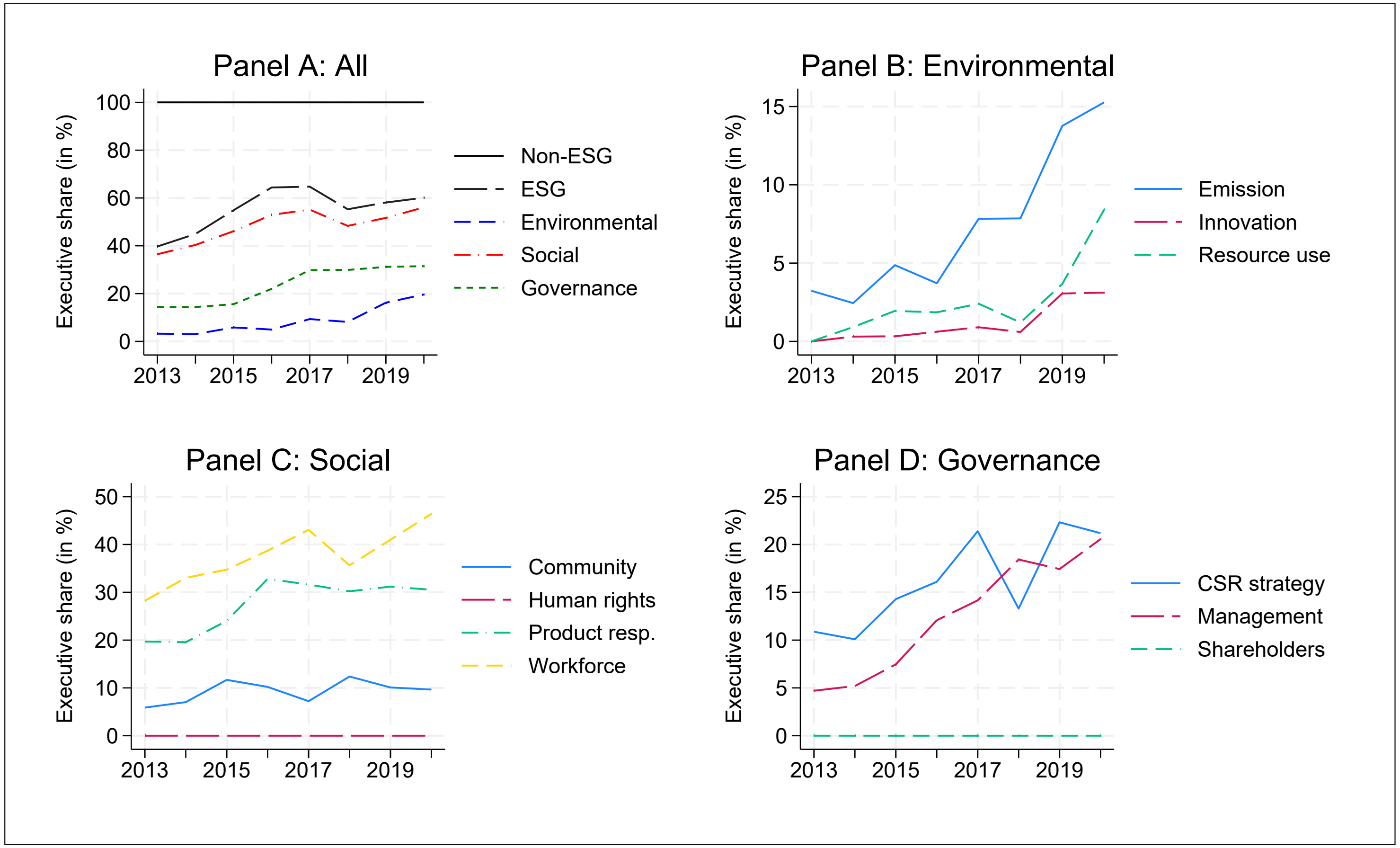


Figure E.6: Share of Executives with at Least one ESG or non-ESG-Related Performance Metric (STI)
The figure shows the share of executives with at least one ESG- or non-ESG performance metric in their short-term incentive contracts (STI). Panel A differentiates between ESG and non-ESG metrics, while also splitting the ESG part into its three main categories: Environmental, Social, and Governance. Panels B to D show the corresponding executive shares for each of the three subcategories in each main ESG category, i.e., Environmental (Panel B), Social (Panel C), and Governance (Panel D). Analyses are based on a sample of 674 executives from 73 firms that have been constituents of the EURO STOXX 50 and the STOXX Europe 50 for at least ten days between 2013 and 2020.

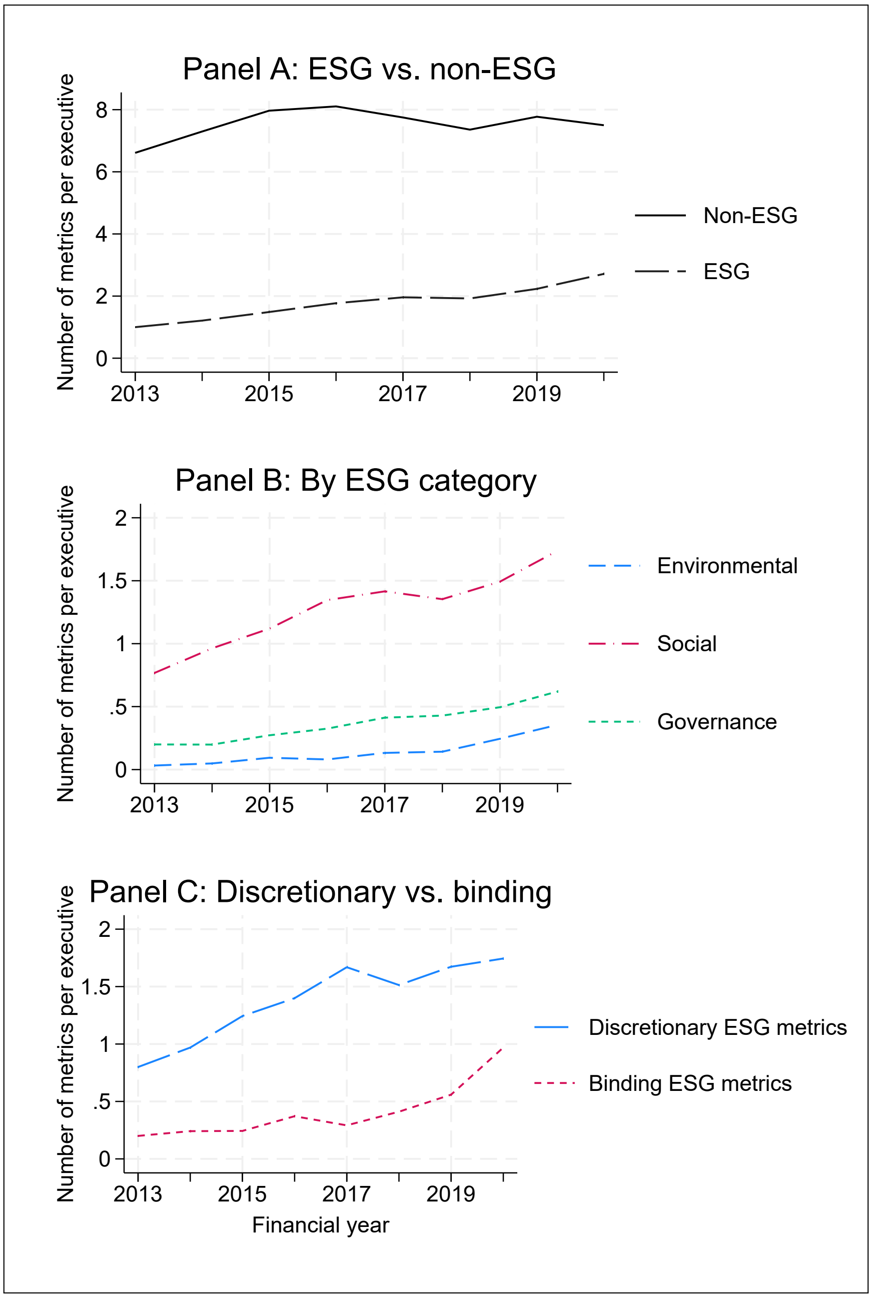


Figure E.7: Number of ESG vs. non-ESG Performance Metrics per Executive (STI)
The figure shows the number of performance metrics in the short-term incentive contract (STI) of the average executive in our sample. Panel A differentiates between ESG and non-ESG metrics and Panel B splits ESG up into its three main categories Environmental, Social, and Governance. Panel C differentiates between discretionary and binding ESG metrics. We define binding performance metrics as metrics where the firm commits already at the beginning of the fiscal year with what weight it will consider each of them in the calculation of realized STI at year-end. For discretionary metrics, however, the respective weights are not known to the executive ex ante. For more information about the design of discretionary and binding performance metrics, see Section 6.4.1.

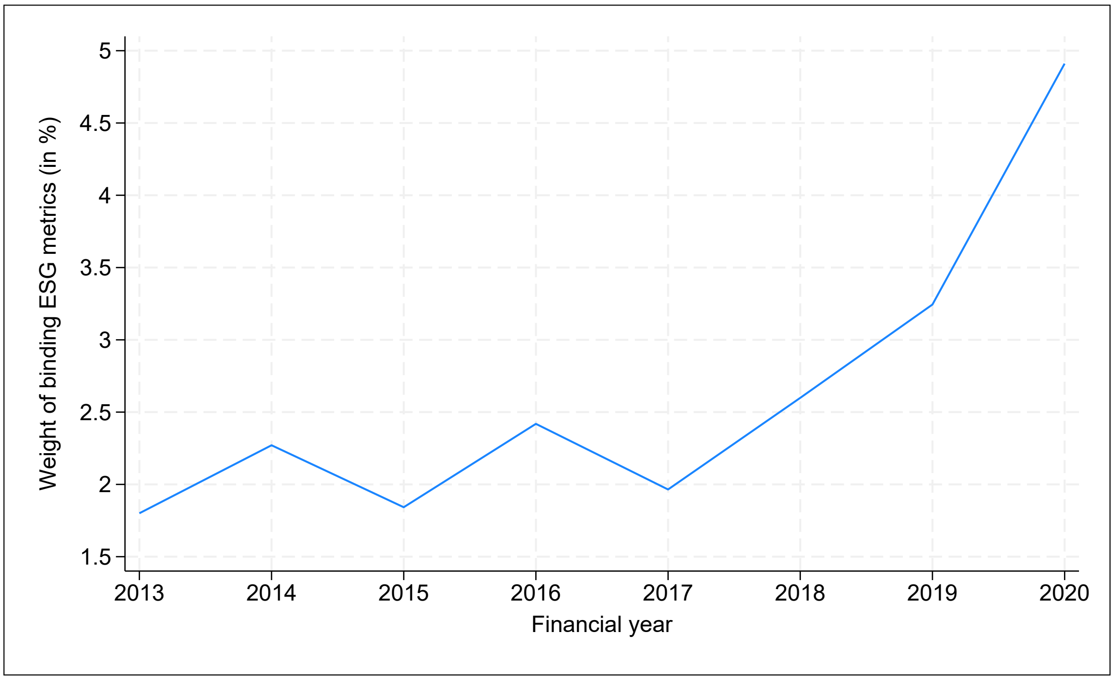


Figure E.8: Weight of Binding ESG Performance Metrics in STI
The figure shows the total weight $w_{i,t}^{B,ESG}$ of all binding ESG metrics together in the short-term incentive contract (STI) of the average executive in our sample. For more information about the design of binding performance metrics and their weights, see Section 6.4.1.

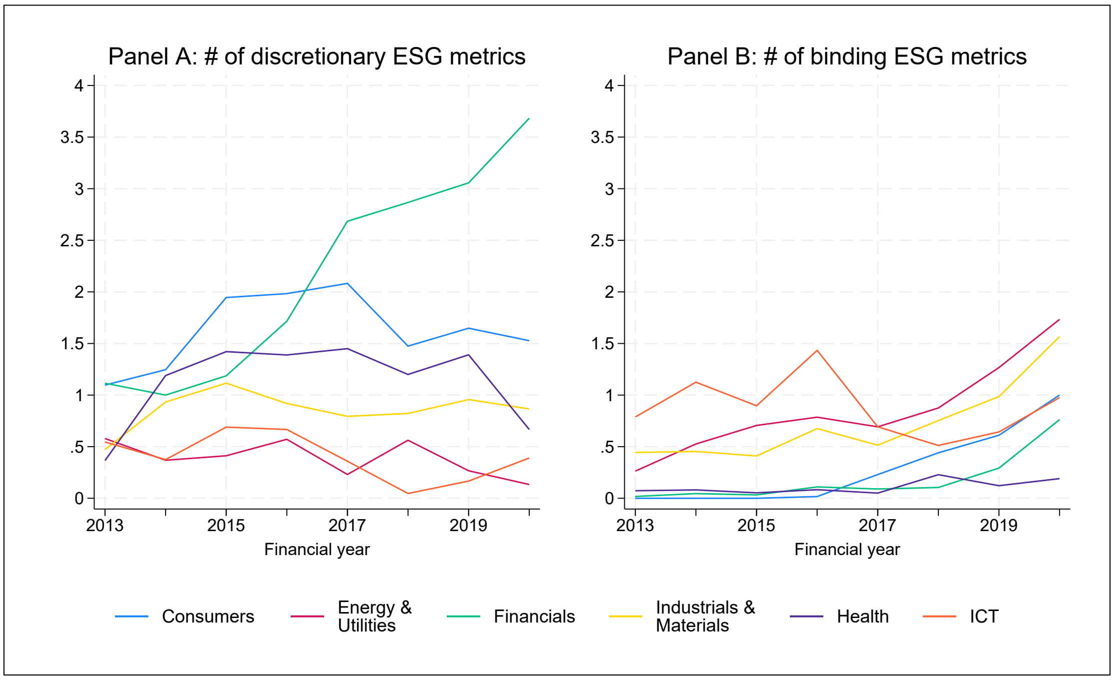


Figure E.9: Number of Binding and Discretionary ESG Performance Metrics in STI by Industry

The figure shows the number of discretionary (Panel A) and binding (Panel B) ESG performance metrics in the short-term incentive contract (STI) of the average executive in a given industry. For more information about the design of discretionary and binding performance metrics, see Section 6.4.1. Industry classification is based on the primary Global Industry Classification Standard (GICS) sector code, which we aggregate to the six major industry sectors Consumers, Energy & Utilities, Financials, Industrials & Materials, Health, and Information & Communication Technology (ICT).

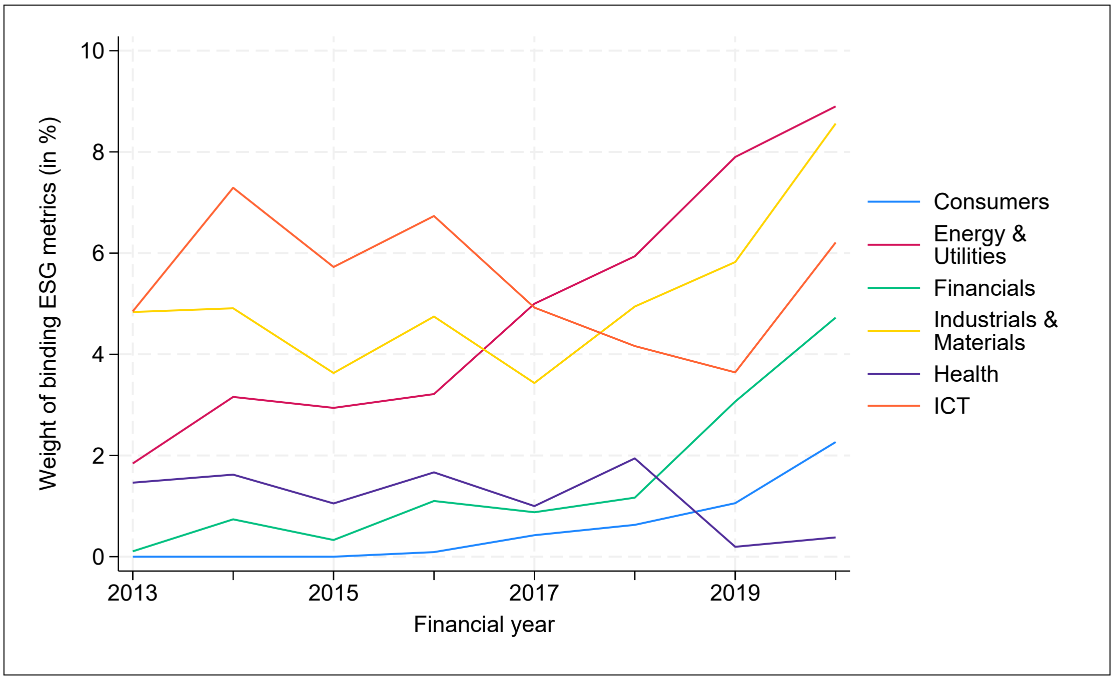


Figure E.10: Weight of Binding ESG Performance Metrics in STI by Industry
The figure shows the sum of the ex-ante weights of all binding ESG metrics in the short-term incentive contract (STI) of the average executive in a given industry. For more information about the design of binding ESG metrics and their weights, see Section 6.4.1. Industry sectors are defined as in Figure E.9.

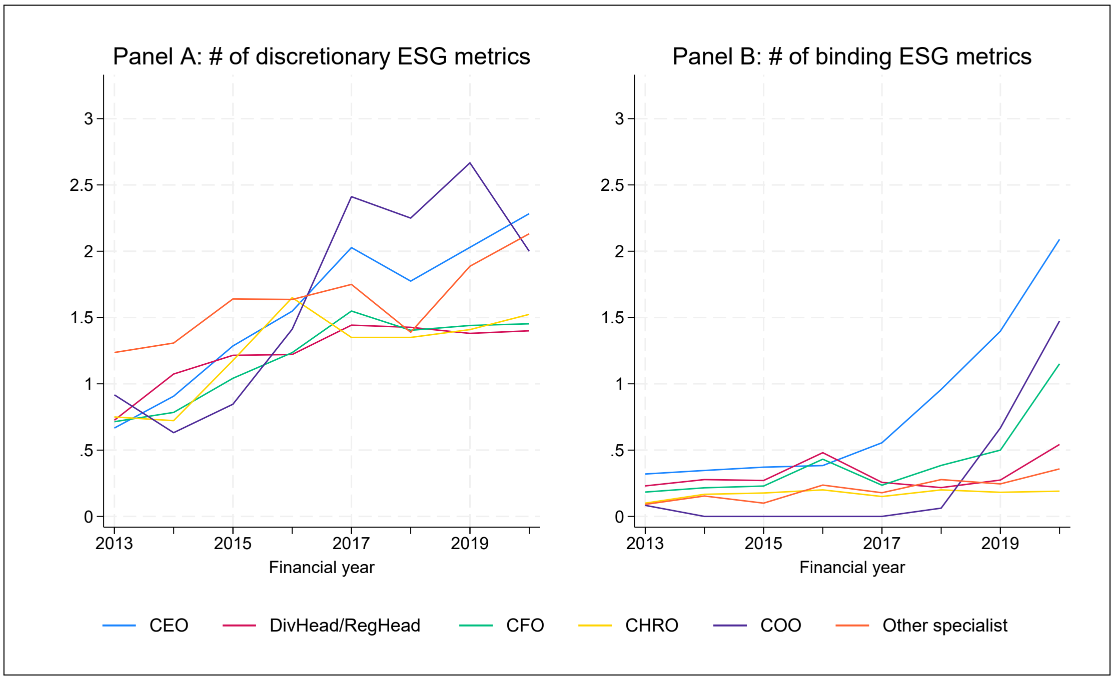


Figure E.11: Number of Binding and Discretionary ESG Performance Metrics in STI by Board Position
The figure shows the number of discretionary (Panel A) and binding (Panel B) ESG performance metrics in the short-term incentive contract (STI) of the average executive in a given board position. We differentiate between the chief executive officer (CEO), chief financial officer (CFO), chief human resources officer (CHRO), and chief operating officer (COO). “DivHead/RegHead” denotes executives who are chairing a division, region, or are responsible for a product segment. All other expert C-suite positions (e.g., chief marketing officer, chief sales officer, chief legal officer, etc.) are grouped as “Other specialist”.

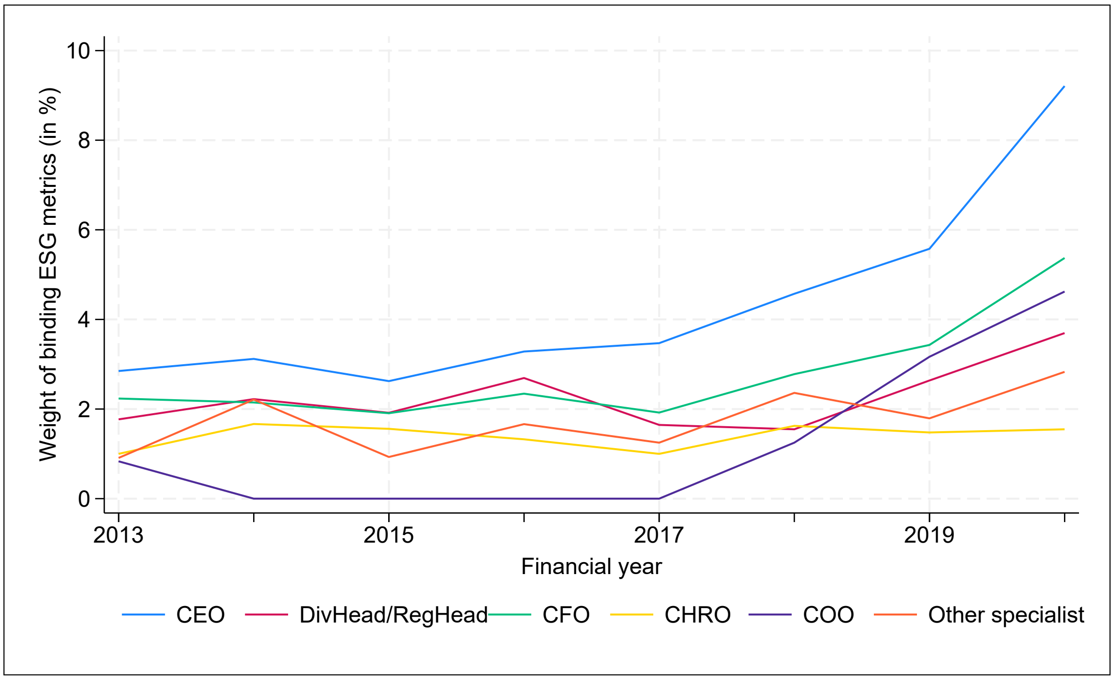


Figure E.12: Weight of Binding ESG Performance Metrics in STI by Board Position
The figure shows the sum of the ex-ante weights of all binding ESG metrics in the short-term incentive contract (STI) of the average executive in a given board position. For more information about the design of binding ESG metrics and their weights, see Section 6.4.1. For a description of the different executive C-suite positions, see Figure E.11.

Table E.3: Firm Sample Composition

This table reports the composition of our sample of the largest European listed firms, differentiated by industry sector (Panel A) and headquarter country (Panel B). For industry sector classification, we apply the primary Global Industry Classification Standard (GICS) sector code from LSEG Data & Analytics, made up of 11 industry sectors and aggregate them to six major industry sectors comprising Consumers, Energy & Utilities, Financials, Industrials & Materials, Health, and Information & Communication Technology (ICT).

	Firm-year observations
Panel A: Industry sector	
Consumers (discretionary & staples)	120
Energy & Utilities	48
Financials	176
Industrials & Materials	112
Health	64
Information & Communication Technology (ICT)	64
Total	584
Panel B: Headquarter country	
Australia	16
Denmark	8
Finland	16
France	104
Germany	136
Ireland	8
Italy	24
Netherlands	24
Spain	40
Switzerland	72
United Kingdom	136
Total	584

Table E.4: Summary Statistics for Firm-Level Variables

Reported are summary statistics on firm characteristics (at the firm level) in a given year. We winsorize all balance sheet items at the 1%-level in each tail of the distribution. For variable definitions, see E.1.

		Obs (1)	Mean (2)	S.D. (3)	P5 (4)	P10 (5)	P25 (6)	P50 (7)	P75 (8)	P90 (9)	P95 (10)
Number of executives	(#)	584	4	3	1	1	2	3	6	9	11
Number of employees	(#)	584	112,332	105,617	18,914	26,267	45,202	87,850	137,335	222,000	308,000
Market capitalization	(10^6€)	584	59,818	45,431	17,437	20,872	29,022	44,923	73,939	117,582	169,950
Total assets	(10^6€)	584	321,626	477,263	15,090	24,191	39,492	87,278	394,412	954,882	1,421,095
EBITDA	(10^6€)	584	10,367	8,035	1,619	2,526	4,318	8,562	13,932	19,503	26,470
EBIT	(10^6€)	584	7,365	5,721	1,242	1,732	3,040	5,706	10,431	14,940	18,671
Net income	(10^6€)	584	3,121	3,453	-1,494	397	1,369	2,581	4,705	7,316	9,183
Book to market ratio	(%)	584	69.21	54.48	9.21	13.79	28.16	53.02	97.86	146.92	172.12
Book leverage	(%)	584	24.94	13.69	2.80	3.93	15.64	24.63	34.72	43.81	49.22
Net PPE/total assets	(%)	584	19.94	20.16	0.24	0.40	1.31	11.17	29.41	46.79	64.18
ROA	(%)	558	5.37	6.43	-0.22	0.33	0.94	4.22	7.79	11.97	14.83
ROE	(%)	583	15.05	17.99	-5.09	0.79	6.41	11.71	18.39	34.52	55.35
Dividends/earnings	(%)	471	52.30	21.78	10.25	22.22	39.88	52.29	68.18	79.92	86.99
Stock return	(%)	584	8.81	22.70	-31.17	-22.19	-4.68	9.26	23.35	37.04	44.12
Stock-to-accounting vola. (hist.)		568	75.61	138.62	3.46	5.13	8.83	21.18	90.53	176.03	397.27
Board independence	(%)	584	71.75	19.84	38.46	45.45	58.33	75.00	87.50	100.00	100.00
Female board membership	(%)	584	31.23	10.25	14.29	17.65	25.00	30.77	38.46	42.86	45.45
Institutional ownership	(%)	584	50.26	15.85	24.79	31.57	38.93	50.88	59.53	70.82	76.48
Block ownership	(%)	584	7.22	16.04	0.00	0.00	0.00	0.00	10.14	26.52	46.57
Emission pledge Y/N	(1/0)	583	1.00	0.06	1.00	1.00	1.00	1.00	1.00	1.00	1.00
CO_2 (Scope 1)	($10^6 t$)	542	6.19	18.17	0.01	0.02	0.03	0.18	1.50	18.59	24.92
CO_2 (Scope 1) (hist.)	($10^6 t$)	568	7.43	22.08	0.01	0.02	0.05	0.16	2.05	18.21	37.78
Log(total assets)		584	25.50	1.43	23.44	23.91	24.40	25.19	26.70	27.58	27.98
Log(book to market ratio)		580	-0.69	0.87	-2.33	-1.95	-1.26	-0.63	-0.02	0.39	0.56
Log(CO_2)		542	12.56	2.62	9.38	9.71	10.37	12.11	14.22	16.74	17.03

Table E.5: Executive Characteristics and Positions

Reported are summary statistics on executive-level variables in a given year. We hand-collect information on an executive's age, gender, tenure, and position within the executive/management board. For the latter, we differentiate between the chief executive officer (CEO), chief financial officer (CFO), chief human resources officer (CHRO), and chief operating officer (COO). All other expert C-suite positions (e.g., chief marketing officer, chief sales officer, chief legal officer, etc.) are grouped as "Other specialists". For executives, who are chairing a division, region, or are responsible for a product segment, we create the job position "Division/regional head". We winsorize age and tenure at the 1%-level in each tail of the distribution.

		Obs (1)	Mean (2)	S.D. (3)	P5 (4)	P25 (5)	P50 (6)	P75 (7)	P95 (8)
Panel A: Executive characteristics									
Age	(years)	2,560	54	5	46	51	54	58	63
Tenure	(years)	2,216	6	5	1	2	4	8	15
Female Y/N	(1/0)	2,560	0.107	0.308	0.000	0.000	0.000	0.000	1.000
Panel B: Executive positions									
CEO		576							
CFO		399							
CHRO		148							
COO		141							
Others		1,264							
Chairman		23							
Divisional Head for Geographic Region or Product Segment		825							
Other specialists		416							
Total		2,528							

Table E.6: Summary Statistics for Short-Term Incentive Pay (STI)

Reported are summary statistics for the main remuneration elements in executives' short-term incentive contracts (STI) at the executive-year level. The superscripts *B* and *D* designate *binding* and *discretionary* performance metrics, respectively. The superscripts *ESG* and *nESG* designate *ESG* and *non-ESG* metrics. For more information about the design of binding and discretionary performance metrics, see Section 6.4. For variable definitions, see E.1.

		Obs (1)	Mean (2)	S.D. (3)	P5 (4)	P10 (5)	P25 (6)	P50 (7)	P75 (8)	P90 (9)	P95 (10)
Target STI amount:											
$TSTI_{i,t}$	(€)	1,915	988,677	577,729	366,600	420,555	623,583	813,934	1,240,250	1,670,918	2,200,000
$TSTI_{i,t}/Base\ salary_{i,t}$	(%)	1,915	100.614	39.665	50	60	80	100	105	150	173
Number of binding STI metrics:											
$\#^{B,ESG}_{i,t}$	(#)	2,609	0.410	1.566	0	0	0	0	0	1	3
$\#^{B,nESG}_{i,t}$	(#)	2,599	4.167	3.983	0	0	2	3	5	9	12
$\#^{B,ESG}_{i,t} + \#^{B,nESG}_{i,t}$	(#)	2,599	4.578	4.337	0	0	2	4	6	10	13
Weight of binding STI metrics:											
$w^{B,ESG}_{i,t}$	(%)	2,599	2.625	7.228	0	0	0	0	0	15	20
$w^{B,nESG}_{i,t}$	(%)	2,599	73.175	34.605	0	0	60	85	100	100	100
$w^{B,ESG}_{i,t} + w^{B,nESG}_{i,t}$	(%)	2,599	75.800	35.462	0	0	65	100	100	100	100
Number of discretionary STI metrics:											
$\#^{D,ESG}_{i,t}$	(#)	2,609	1.374	2.375	0	0	0	0	2	5	5
$\#^{D,nESG}_{i,t}$	(#)	2,599	3.370	5.787	0	0	0	0	5	11	17
$\#^{D,ESG}_{i,t} + \#^{D,nESG}_{i,t}$	(#)	2,599	4.750	7.473	0	0	0	0	7	16	21
Weight of discretionary STI metrics:											
$w^{D}_{i,t}$	(%)	2,599	24.195	35.460	0	0	0	0	35	100	100
$STI\ hurdle\ Y/N_{i,t}$	(1/0)	2,677	0.631	0.483	0	0	0	1	1	1	1
$STI\ cap\ Y/N_{i,t}$	(1/0)	2,677	0.959	0.199	1	1	1	1	1	1	1
$STI\ cap_{i,t}/Base\ salary_{i,t}$	(%)	2,243	176.545	70.793	50	98	130	180	200	250	281
$STI\ board\ discretion\ allowed\ Y/N_{i,t}$	(1/0)	2,666	0.728	0.445	0	0	0	1	1	1	1
$STI\ ex\ post\ multiplier\ (ESG)\ Y/N_{i,t}$	(1/0)	2,677	0.040	0.195	0	0	0	0	0	0	0
$STI\ ex\ post\ multiplier\ (nESG)\ Y/N_{i,t}$	(1/0)	2,677	0.142	0.349	0	0	0	0	0	1	1

Table E.7: Summary Statistics for Long-Term Incentive Pay (LTI)

Reported are summary statistics for the main remuneration elements in executives' long-term incentive contracts (LTI) at the executive-year level. Information on the numbers and weights of metrics in LTI plans always refer to the first year of granting. The superscripts B and D designate *binding* and *discretionary* performance metrics, respectively. The superscripts ESG and $nESG$ designate ESG and *non-ESG* metrics. For more information about the design of binding and discretionary performance metrics, see Section 6.4. For variable definitions, see E.1.

		Obs (1)	Mean (2)	S.D. (3)	P5 (4)	P10 (5)	P25 (6)	P50 (7)	P75 (8)	P90 (9)	P95 (10)
Number of binding LTI metrics:											
$\#_{i,t}^{B,ESG}$	(#)	2,499	0.201	0.716	0	0	0	0	0	0	2
$\#_{i,t}^{B,nESG}$	(#)	2,443	2.574	1.890	0	1	2	2	3	5	5
$\#_{i,t}^{B,ESG} + \#_{i,t}^{B,nESG}$	(#)	2,443	2.780	2.120	0	1	2	2	3	5	6
Weights of binding LTI metrics:											
$w_{i,t}^{B,ESG}$	(%)	2,443	2.479	9.111	0	0	0	0	0	0	20
$w_{i,t}^{B,nESG}$	(%)	2,443	89.169	27.119	0	60	100	100	100	100	100
$w_{i,t}^{B,ESG} + w_{i,t}^{B,nESG}$	(%)	2,443	91.649	26.329	0	70	100	100	100	100	100
Number of discretionary LTI metrics:											
$\#_{i,t}^{D,ESG}$	(#)	2,499	0.188	0.672	0	0	0	0	0	1	1
$\#_{i,t}^{D,nESG}$	(#)	2,443	0.605	2.041	0	0	0	0	0	3	5
$\#_{i,t}^{D,ESG} + \#_{i,t}^{D,nESG}$	(#)	2,443	0.797	2.601	0	0	0	0	0	4	7
Weight of discretionary LTI metrics:											
$w_{i,t}^{D}$	(%)	2,443	8.351	26.329	0	0	0	0	0	30	100
$LTI\ hurdle\ Y/N_{i,t}$	(1/0)	2,677	0.759	0.428	0	0	1	1	1	1	1
$LTI\ cap\ Y/N_{i,t}$	(1/0)	2,677	0.410	0.492	0	0	0	0	1	1	1
$LTI\ cap_{i,t}/Base\ salary_{i,t}$	(%)	1,097	292.521	119.171	133	140	200	275	380	438	500
$LTI\ board\ discretion\ allowed\ Y/N_{i,t}$	(1/0)	2,276	0.753	0.431	0	0	1	1	1	1	1
$LTI\ ex\ post\ multiplier\ (ESG)\ Y/N_{i,t}$	(1/0)	2,677	0.000	0.000	0	0	0	0	0	0	0
$LTI\ ex\ post\ multiplier\ (nESG)\ Y/N_{i,t}$	(1/0)	2,677	0.016	0.124	0	0	0	0	0	0	0

Table E.8: STI and LTI Performance Metrics by ESG Category

Reported are summary statistics for the number of ESG and non-ESG performance metrics by different categories. In the case of ESG metrics, we show the number of metrics for each of the three categories Environmental, Social, and Governance. In the case of non-ESG metrics, we distinguish between financial and non-financial metrics. We do not distinguish between binding and discretionary metrics in this table. Panel A focuses on short-term incentive contracts (STI), whereas Panel B reports numbers for long-term incentive contracts (LTI).

	Obs (1)	Mean (2)	S.D. (3)	P5 (4)	P10 (5)	P25 (6)	P50 (7)	P75 (8)	P90 (9)	P95 (10)
Panel A: STI										
Number of ESG metrics:	2,609	1.784	2.762	0	0	0	1	3	5	6
$\#(Environmental)_{i,t}$	2,609	0.141	0.601	0	0	0	0	0	0	1
$\#(Social)_{i,t}$	2,609	1.273	1.988	0	0	0	0	2	4	5
$\#(Governance)_{i,t}$	2,609	0.369	0.772	0	0	0	0	1	1	2
Number of non-ESG metrics:	2,599	7.537	6.552	1	2	3	5	9	16	22
$\#(Financial)_{i,t}$	2,599	3.595	2.880	1	1	2	3	4	6	8
$\#(Non\text{-}financial)_{i,t}$	2,599	3.942	4.722	0	0	1	2	5	10	14
Panel B: LTI										
Number of ESG metrics:	2,514	0.399	0.963	0	0	0	0	0	2	2
$\#(Environmental)_{i,t}$	2,514	0.058	0.360	0	0	0	0	0	0	0
$\#(Social)_{i,t}$	2,514	0.261	0.721	0	0	0	0	0	1	2
$\#(Governance)_{i,t}$	2,514	0.070	0.298	0	0	0	0	0	0	1
Number of non-ESG metrics:	2,443	3.179	2.319	1	1	2	3	4	5	7
$\#(Financial)_{i,t}$	2,443	2.503	1.266	1	1	2	2	3	4	5
$\#(Non\text{-}financial)_{i,t}$	2,443	0.676	1.699	0	0	0	0	0	2	4

Table E.9: STI Target Achievement and Realized Pay

Panel A shows summary statistics for target fulfillment rates (achievement rates) in short-term incentive pay (STI). Panel B shows summary statistics for realized compensation (base salary, STI, and LTI), board discretion, and deferred pay. For more information about the design of STI and LTI, see Section 6.4. For variable definitions, see E.1.

		Obs (1)	Mean (2)	S.D. (3)	P5 (4)	P10 (5)	P25 (6)	P50 (7)	P75 (8)	P90 (9)	P95 (10)
Panel A: Target fulfillment rates in STI											
Overall target fulfillment rate:											
$f_{i,t}$	(%)	1,860	104	40	36	59	86	102	123	145	166
Fulfillment rate of binding targets:											
$f_{i,t}^{B,ESG}$	(%)	231	103	31	44	69	92	100	121	148	150
$f_{i,t}^{B,nESG}$	(%)	1,324	100	40	34	51	82	100	119	139	159
$f_{i,t}^{B}$	(%)	1,585	101	39	37	54	82	100	120	143	159
Fulfillment rate of discretionary targets:											
$f_{i,t}^{D}$	(%)	714	108	41	27	63	89	104	124	160	180
Panel B: Realized pay (base, STI, LTI)											
$RSTI_{i,t}/Base\ salary_{i,t}$	(%)	2,388	84	54	0	19	47	79	115	149	170
$RSTI_{i,t}$	(€)	2,388	821,716	612,654	0	174,867	424,817	707,572	1,053,200	1,608,147	2,027,285
$Base\ salary_{i,t}$	(€)	2,464	1,093,465	576,840	513,338	600,000	720,000	882,466	1,300,000	1,923,100	2,339,052
$RSTI_{i,t} + Base\ salary_{i,t}$	(€)	2,560	1,819,335	985,311	511,539	841,699	1,207,902	1,622,384	2,349,000	3,053,200	3,621,199
$RLTI_{i,t}$	(€)	2,508	1,726,691	1,334,493	30,000	303,000	855,795	1,423,006	2,229,780	3,627,090	4,412,120
$RSTI_{i,t} + Base\ salary_{i,t} + RLTI_{i,t}$	(€)	2,575	3,515,619	2,026,066	942,000	1,477,724	2,165,905	3,008,665	4,574,020	6,397,592	7,565,600
$Board\ exercised\ discretion\ Y/N_{i,t}$	(1/0)	2,609	0.13	0.34	0.00	0.00	0.00	0.00	0.00	1.00	1.00
$STI\ partly\ deferred\ Y/N_{i,t}$	(1/0)	2,677	0.51	0.50	0.00	0.00	0.00	1.00	1.00	1.00	1.00
$STI\ deferral_{i,t}/RSTI_{i,t}$	(%)	947	53.40	17.43	33.30	33.33	50.00	50.00	60.00	80.00	80.00
$STI\ deferral\ period_{i,t}$	(years)	1,034	3.25	0.82	2.00	2.00	3.00	3.00	4.00	5.00	5.00

Table E.10: STI Variance Decomposition

This table shows the decomposition of the variance of executives' realized short-term incentive compensation $RSTI_{i,t}$ into its different components. Column 2 shows total STI variance $Var\,(RSTI_{i,t})$. In Panel A, total variance includes both within-executive and between-executive STI variation. Panel B only considers within-executive variation. In each panel, we estimate the variance decomposition in the largest possible sample (*All observations*), in the subsample of observations without board discretion (*No discretion*), and in the subsample of observations with at least one ESG metric (*Min. 1 ESG metric*). The variance shares reported in columns 4, 6, 8, and 9 to 11 correspond to the variance and covariance terms in Equation 6.5 divided by total STI variance. For example, the variance share of binding ESG metrics in column 4 is defined as $Var(RSTI_{i,t}^{B,ESG})/Var(RSTI_{i,t})$. Columns 3, 5, and 7 report the ex ante weights that firms assign to the different STI performance metrics, and are calculated as the sample averages of $w_{i,t}^{B,ESG}$, $w_{i,t}^{B,nESG}$, and $w_{i,t}^{D}$. As in previous tables, superscripts B and D designate binding and discretionary performance metrics, and superscripts ESG and $nESG$ designate ESG and non-ESG metrics. For more information on the variance decomposition approach, see Section 6.6.1.2.

			Binding metrics				Discretionary metrics		Covariance terms (2×cov)		
			$RSTI_{i,t}^{B,ESG}$		$RSTI_{i,t}^{B,nESG}$		$RSTI_{i,t}^{D}$		$(RSTI_{i,t}^{B,ESG}, RSTI_{i,t}^{B,nESG})$	$(RSTI_{i,t}^{B,ESG}, RSTI_{i,t}^{D})$	$(RSTI_{i,t}^{B,nESG}, RSTI_{i,t}^{D})$
	obs	$Var(RSTI_{i,t})$ $\times 10^{-6}$	STI weight	Var share	STI weight	Var share	STI weight	Var share	Var share	Var share	Var share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: Within- and between-executive STI variance											
All observations	1,076	492,320	0.035	0.019	0.875	0.881	0.090	0.093	0.044	−0.004	−0.034
No discretion	926	505,566	0.031	0.017	0.884	0.898	0.085	0.086	0.043	−0.006	−0.038
Min. 1 ESG metric	443	538,548	0.085	0.035	0.697	0.671	0.219	0.145	0.144	−0.053	0.058
Panel B: Within-executive STI variance											
All observations	1,074	281,694	0.035	0.010	0.875	0.872	0.090	0.077	0.024	−0.004	0.021
No discretion	924	275,669	0.031	0.009	0.884	0.913	0.085	0.069	0.019	−0.007	−0.003
Min. 1 ESG metric	443	261,754	0.085	0.025	0.697	0.649	0.219	0.152	0.064	−0.013	0.123

Table E.11: Annual Variation in Overall Target Achievement by CEO

We report cross-sectional OLS regressions for the annual variation in target fulfillment of CEOs. We calculate the dependent variable as the standard deviation of $f_{i,t}$ between 2013 and 2020. A high value means that overall performance achievement, measured across all the different performance metrics in the STI plan of CEO i, fluctuates more between 2013 and 2020. The dependent variable is regressed on the numbers and weights of ESG and non-ESG metrics in the STI contract, as well as on various executive and firm characteristics and industry fixed effects. These independent variables are calculated as time-series averages over our sample period. For example, we calculate the *Number of ESG metrics*$_i$ as the average number of ESG metrics of executive i between 2013 and 2020. Robust standard errors are reported in parentheses. $^{*}p < 0.1;^{**}p < 0.05;^{***}p < 0.01$.

	Dependent variable: *SD(Target fulfillment f)*$_i$					
	(1)	(2)	(3)	(4)	(5)	(6)
Number of ESG metrics$_i$ ($\#^{B,ESG} + \#^{D,ESG}$)	-0.959**	-0.933**	-1.143*			
	(0.466)	(0.458)	(0.629)			
Number of non-ESG metrics$_i$ ($\#^{B,nESG} + \#^{D,nESG}$)	-0.089	-0.033	0.143			
	(0.572)	(0.566)	(0.604)			
Weight of binding ESG metrics$_i$ ($w^{B,ESG}$)				-0.622**	-0.748**	-0.519*
				(0.298)	(0.346)	(0.301)
Weight of binding non-ESG metrics$_i$ ($w^{B,nESG}$)				-0.071	-0.135	-0.149
				(0.134)	(0.132)	(0.121)
Log(Base salary)$_i$		-3.409	-9.979		-8.815	-13.045
		(8.205)	(10.534)		(8.492)	(10.539)
Tenure$_i$		-0.613	-0.549		-0.854	-0.796
		(1.104)	(0.881)		(0.992)	(0.829)
Age$_i$		0.397	-0.073		0.0648	0.058
		(0.796)	(0.643)		(0.806)	(0.613)
Female$_i$		-3.596	-18.831		-2.712	-14.700
		(9.608)	(16.993)		(9.359)	(16.074)
Log(total assets)$_i$		-0.926	-0.306		-1.609	-1.388
		(1.954)	(3.282)		(1.941)	(3.112)
Industry F.E.	No	No	Yes	No	No	Yes
Observations	65	65	65	65	65	65
Adjusted R^2	-0.01	-0.08	0.00	0.02	-0.02	0.02

Table E.12: STI Variance Secomposition by Industry

This table shows the decomposition of the variance of executives' realized short-term incentive compensation $RSTI_{i,t}$ for different industry sectors. Column 2 shows total STI variance $Var\,(RSTI_{i,t})$. In Panel A, total variance includes both within-executive and between-executive STI variation. Panel B only considers within-executive variation. The variance shares reported in columns 4, 6, 8, and 9 to 11 correspond to the variance and covariance terms in Equation 6.5 divided by total STI variance. For example, the variance share of binding ESG metrics in column 4 is defined as $Var(RSTI_{i,t}^{B,ESG})/Var(RSTI_{i,t})$. Columns 3, 5, and 7 report the ex ante weights that firms assign to the different STI performance metrics, and are calculated as the sample averages of $w_{i,t}^{B,ESG}$, $w_{i,t}^{B,nESG}$, and $w_{i,t}^{D}$. As in previous tables, superscripts B and D designate binding and discretionary performance metrics, and superscripts ESG and $nESG$ designate ESG and non-ESG metrics. For more information on the variance decomposition approach, see Section 6.6.1.2.

			Binding metrics				Discretionary metrics		Covariance terms (2×cov)		
			$RSTI_{i,t}^{B,ESG}$		$RSTI_{i,t}^{B,nESG}$		$RSTI_{i,t}^{D}$		$(RSTI_{i,t}^{B,ESG}, RSTI_{i,t}^{B,nESG})$	$(RSTI_{i,t}^{B,ESG}, RSTI_{i,t}^{D})$	$(RSTI_{i,t}^{B,nESG}, RSTI_{i,t}^{D})$
	obs	$Var(RSTI_{i,t})$ $\times 10^{-6}$	STI weight	Var share	STI weight	Var share	STI weight	Var share	Var share	Var share	Var share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: Within- and between-executive STI variance											
Consumers	215	801,191	0.002	0.000	0.909	1.026	0.090	0.112	−0.001	−0.001	−0.137
Energy & Utilities	74	414,271	0.046	0.061	0.885	0.619	0.069	0.073	0.271	−0.023	−0.001
Financials	217	340,937	0.012	0.009	0.897	0.743	0.091	0.065	0.068	0.014	0.100
Industrials & Materials	288	394,729	0.064	0.023	0.806	0.882	0.130	0.126	−0.017	−0.011	−0.003
Health	116	252,268	0.000	0.000	0.936	1.012	0.064	0.174	0.000	0.000	−0.187
ICT	166	483,301	0.075	0.051	0.877	0.752	0.048	0.031	0.187	−0.008	−0.012
Panel B: Within-executive STI variance											
Consumers	214	622,787	0.002	0.000	0.908	0.955	0.090	0.042	0.003	−0.001	0.001
Energy & Utilities	74	226,577	0.046	0.029	0.885	0.625	0.069	0.086	0.142	0.000	0.119
Financials	216	153,286	0.012	0.005	0.896	0.871	0.091	0.082	0.031	−0.008	0.019
Industrials & Materials	288	277,063	0.064	0.017	0.806	0.769	0.130	0.120	0.036	−0.008	0.066
Health	116	163,679	0.000	0.000	0.936	0.952	0.064	0.157	0.000	0.000	−0.109
ICT	166	131,921	0.075	0.043	0.877	0.864	0.048	0.051	0.032	−0.006	0.017

Table E.13: ESG Pay and Firm and Employee Characteristics

We report panel regressions for various measures of ESG adoption in executive STI as function of firm and executive characteristics as well as fixed effects for industry sectors, executive positions, and years. The dependent variables are defined as follows. Column 1: At least 1 $ESG_D = \mathbb{1}(\#_{i,t}^{D,ESG} \geq 1)$; column 2: Log(1+No. ESG_D) $= Ln(1 + \#_{i,t}^{D,ESG})$; column 3: No. ESG_D/No. all metrics $= \#_{i,t}^{D,ESG}/(\#_{i,t}^{D} + \#_{i,t}^{B})$; column 4: Weight All $D = w_{i,t}^{D}$; column 5: At least 1 $ESG_B = \mathbb{1}(\#_{i,t}^{B,ESG} \geq 1)$; column 6: Log(1+No. ESG_B) $= Ln(1 + \#_{i,t}^{B,ESG})$; column 7: No. ESG_B/No. all metrics $= \#_{i,t}^{B,ESG}/(\#_{i,t}^{D} + \#_{i,t}^{B})$; column 8: Weight $ESG_B = w_{i,t}^{B,ESG}$. The independent variables are defined in E.1. Robust standard errors are clustered by firm and reported in parentheses. $^{*}p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

	Discretionary ESG metrics (ESG_D)				Binding ESG metrics (ESG_B)			
	At least 1 ESG_D (1)	Log(1+No. ESG_D) (2)	No. ESG_D / No. all metrics (3)	Weight All D (4)	At least 1 ESG_B (5)	Log(1+No. ESG_B) (6)	No. ESG_B / No. all metrics (7)	Weight ESG_B (8)
Historical Log(Avg. CO_2)	-0.069	-0.132**	-0.027*	-3.619	-0.019	-0.031	-0.003	-0.084
	(0.044)	(0.064)	(0.015)	(2.881)	(0.024)	(0.036)	(0.010)	(0.496)
Historical stock-to-accounting volatility ($\times 10^{-2}$)	0.024	-0.028	0.007	0.878	0.033**	0.039**	0.020***	0.744*
	(0.029)	(0.044)	(0.005)	(1.771)	(0.013)	(0.018)	(0.007)	(0.388)
Log(total assets)	0.118	0.228**	0.049*	12.732**	-0.042	-0.017	-0.011	-1.028
	(0.072)	(0.113)	(0.022)	(5.153)	(0.064)	(0.093)	(0.026)	(1.304)
Log(book to market ratio)	0.124**	0.216**	0.049**	5.930	0.048	0.036	0.009	0.892
	(0.060)	(0.085)	(0.022)	(4.687)	(0.053)	(0.080)	(0.023)	(1.244)
ROA	-0.001	0.002	-0.001	-0.228	-0.000	0.000	0.000	-0.021
	(0.008)	(0.010)	(0.002)	(0.511)	(0.004)	(0.005)	(0.001)	(0.082)
Stock return	0.002*	0.003*	0.001**	0.133	0.001	0.000	0.000	0.018
	(0.001)	(0.002)	(0.000)	(0.086)	(0.001)	(0.001)	(0.000)	(0.014)
Book leverage	0.000	-0.001	-0.000	-0.139	0.005	0.004	0.002	0.058
	(0.004)	(0.005)	(0.001)	(0.236)	(0.003)	(0.005)	(0.001)	(0.070)
Net PPE/total assets	0.006**	0.011**	0.003**	0.267	-0.000	0.002	0.000	0.041
	(0.003)	(0.005)	(0.001)	(0.201)	(0.002)	(0.003)	(0.001)	(0.047)
DPS	-0.000	0.011	-0.002	0.934	0.055**	0.075*	0.024**	1.063*
	(0.035)	(0.045)	(0.010)	(2.344)	(0.024)	(0.038)	(0.011)	(0.544)
Emissions policy	0.735***	0.911***	0.147***	27.800***	0.242***	0.202**	0.060**	3.449***
	(0.210)	(0.254)	(0.044)	(7.996)	(0.070)	(0.087)	(0.024)	(1.262)
ESG rating	0.033	0.070*	0.004	-1.285	0.002	0.021	0.003	0.371
	(0.028)	(0.039)	(0.008)	(1.824)	(0.019)	(0.025)	(0.006)	(0.336)
Institutional ownership	0.006**	0.010***	0.002**	0.182	-0.006**	-0.004*	-0.001*	-0.051*
	(0.003)	(0.003)	(0.001)	(0.188)	(0.002)	(0.002)	(0.000)	(0.029)
Block ownership	-0.008**	-0.009*	-0.002	-0.445	0.003*	0.002	0.000	0.016
	(0.003)	(0.005)	(0.002)	(0.283)	(0.002)	(0.002)	(0.001)	(0.034)
Board independence	0.003*	0.007**	0.001	0.189*	-0.001	-0.003	-0.001	-0.038
	(0.002)	(0.003)	(0.001)	(0.102)	(0.001)	(0.002)	(0.001)	(0.036)
Female board membership	-0.002	0.002	0.001	-0.045	0.002	0.004	0.001	0.037
	(0.005)	(0.007)	(0.001)	(0.379)	(0.003)	(0.004)	(0.001)	(0.051)
Tenure	-0.001	0.000	0.001	0.049	-0.002	-0.000	0.000	-0.016
	(0.005)	(0.008)	(0.002)	(0.402)	(0.003)	(0.003)	(0.001)	(0.055)
Age	0.004	0.003	-0.000	0.102	0.001	0.001	0.001	0.082**
	(0.003)	(0.004)	(0.001)	(0.218)	(0.002)	(0.003)	(0.001)	(0.039)
Female	-0.082**	-0.089*	-0.009	-2.542	-0.007	-0.006	0.004	0.218
	(0.033)	(0.046)	(0.011)	(1.967)	(0.022)	(0.033)	(0.010)	(0.510)
Industry F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Executive position F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,056	2,056	2,055	2,055	2,056	2,056	2,055	2,055
Adjusted R^2	0.23	0.29	0.26	0.17	0.17	0.16	0.16	0.15

Table E.14: STI Variance Decomposition by Executive Position

This table shows the decomposition of the variance of executives' realized short-term incentive compensation $RSTI_{i,t}$ for different executive positions. Column 2 shows total STI variance $Var\,(RSTI_{i,t})$. In Panel A, total variance includes both within-executive and between-executive STI variation. Panel B only considers within-executive variation. The variance shares reported in columns 4, 6, 8, and 9 to 11 correspond to the variance and covariance terms in Equation 6.5 divided by total STI variance. For example, the variance share of binding ESG metrics in column 4 is defined as $Var(RSTI_{i,t}^{B,ESG})/Var(RSTI_{i,t})$. Columns 3, 5, and 7 report the ex ante weights that firms assign to the different STI performance metrics, and are calculated as the sample averages of $w_{i,t}^{B,ESG}$, $w_{i,t}^{B,nESG}$, and $w_{i,t}^{D}$. As in previous tables, superscripts B and D designate binding and discretionary performance metrics, and superscripts ESG and $nESG$ designate ESG and non-ESG metrics. For more information on the variance decomposition approach, see Section 6.6.1.2.

			Binding metrics				Discretionary metrics		Covariance terms (2×cov)		
			$RSTI_{i,t}^{B,ESG}$		$RSTI_{i,t}^{B,nESG}$		$RSTI_{i,t}^{D}$		$(RSTI_{i,t}^{B,ESG}, RSTI_{i,t}^{B,nESG})$	$(RSTI_{i,t}^{B,ESG}, RSTI_{i,t}^{D})$	$(RSTI_{i,t}^{B,nESG}, RSTI_{i,t}^{D})$
	obs	$Var(RSTI_{i,t})$ $\times 10^{-6}$	STI weight	Var share	STI weight	Var share	STI weight	Var share	Var share	Var share	Var share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: Within- and between-executive STI variance											
CEO	297	760,444	0.043	0.027	0.846	1.005	0.111	0.107	0.020	−0.014	−0.145
CFO	200	213,585	0.038	0.039	0.870	0.867	0.091	0.183	0.036	−0.009	−0.117
CHRO	69	134,169	0.020	0.008	0.915	1.138	0.065	0.141	−0.042	−0.008	−0.236
COO	55	181,326	0.011	0.029	0.886	0.947	0.103	0.057	0.120	−0.006	−0.148
Division/Region Head	318	160,129	0.031	0.012	0.890	1.002	0.079	0.166	−0.052	−0.012	−0.115
Other specialist	124	159,487	0.037	0.013	0.884	1.059	0.079	0.179	−0.030	−0.013	−0.208
Panel B: Within-executive STI variance											
CEO	296	429,993	0.043	0.014	0.845	0.963	0.112	0.095	−0.003	−0.011	−0.059
CFO	199	166,423	0.038	0.015	0.870	0.867	0.092	0.100	0.057	−0.008	−0.031
CHRO	69	80,902	0.020	0.003	0.915	0.903	0.065	0.120	−0.004	0.001	−0.022
COO	55	111,287	0.011	0.008	0.886	0.854	0.103	0.041	0.037	−0.001	0.061
Division/Region Head	318	85,802	0.031	0.010	0.890	0.822	0.079	0.154	0.003	−0.001	0.012
Other specialist	124	117,325	0.037	0.017	0.884	0.926	0.079	0.088	0.098	0.003	−0.133

Table E.15: Tailoring of ESG Metrics to Executive Positions

Reported are panel regressions for the adoption of different ESG performance metrics in the STI contracts of CTOs, CHROs, and CEOs. The dependent variables are binary indicators that equal one if the STI of executive i in year t is linked to a given type of ESG metric: environmental metrics in general in columns 1 and 2 of Panel A; metrics measuring emissions in columns 3 and 4 of Panel A; social metrics in general in columns 1 and 2 of Panel B; workforce-related metrics in columns 3 and 4 of Panel B. The independent variables *CTO dummy* and *CHRO dummy* are binary indicators that equal one if executive i is a CTO or a CHRO, respectively (and zero otherwise). The omitted base category consists of CEOs. Firm, year, and interacted firm-year fixed effects are included as indicated. Robust standard errors are reported in parentheses. $^{*}p < 0.1$;$^{**}p < 0.05$;$^{***}p < 0.01$.

	(1)	(2)	(3)	(4)
Panel A: Environmental metrics				
	Environm. Y/N		Emissions Y/N	
CTO dummy	-0.068***	-0.066**	-0.053***	-0.052**
	(0.023)	(0.026)	(0.020)	(0.024)
CHRO dummy	-0.037*	-0.040**	-0.034*	-0.034**
	(0.020)	(0.017)	(0.018)	(0.016)
Firm F.E. and Year F.E.	Yes	No	Yes	No
Firm-year F.E.	No	Yes	No	Yes
Observations (CEO)	571	209	571	209
Observations (CTO)	92	92	92	92
Observations (CHRO)	158	155	158	155
Adjusted R^2	0.51	0.73	0.53	0.75
Panel B: Social metrics				
	Social Y/N		Workforce Y/N	
CTO dummy	-0.039	-0.065**	-0.050	-0.066**
	(0.039)	(0.028)	(0.039)	(0.027)
CHRO dummy	-0.032	-0.025	-0.004	-0.012
	(0.028)	(0.017)	(0.029)	(0.018)
Firm F.E. and Year F.E.	Yes	No	Yes	No
Firm-year F.E.	No	Yes	No	Yes
Observations (CEO)	571	209	571	209
Observations (CTO)	92	92	92	92
Observations (CHRO)	158	155	158	155
Adjusted R^2	0.61	0.91	0.59	0.89

E.6 Additional Empirical Results

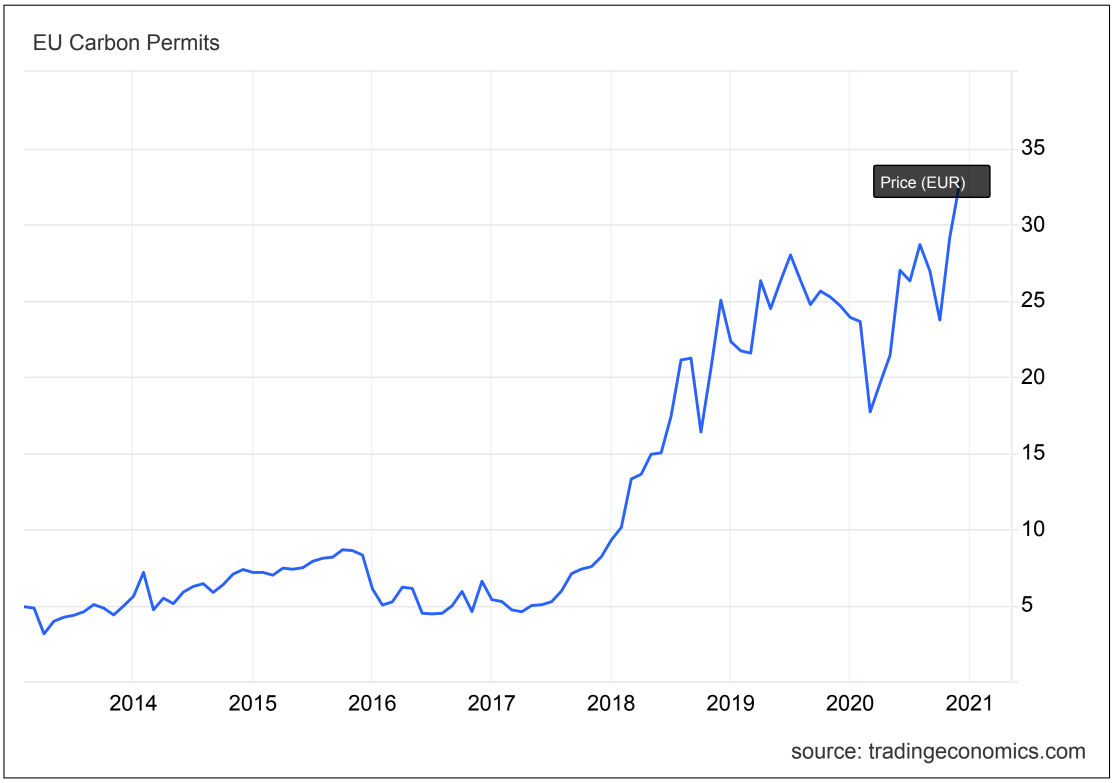


Figure E.13: Evolution of EU Carbon Permit Prices
The figure illustrates the evolution of EU carbon permit prices (in Euros) from 2013 to 2021. Data were downloaded from https://tradingeconomics.com/commodity/carbon on July 9, 2024.

E.7 Overview Firms in Sample – Detailed List

Table E.16: Sample Firms

Firm	Country	Sample period (FY)	STOXX Europe 50	EURO STOXX 50	Comments
ABB	CH	2013-2020	since 09/2008	-	-
Adidas	DE	2013-2020	since 09/2020	since 09/2016	-
Adyen	NL	2018-2020	since 09/2020	since 09/2020	First annual report published in 2018
Ahold Delhaize	NL	2016-2020	-	since 09/2016	Merger in 07/2016 of “Koninklijke Ahold” (NL) and “Delhaize Group” (BE)
Air Liquide	FR	2013-2020	since 09/2011	since 08/1998	-
Airbus	NL	2013-2020	since 09/2017	since 03/2013	-
Allianz	DE	2013-2020	since 08/1998	since 08/1998	-
Amadeus IT	ES	2013-2020	-	since 09/2018	-
Anheuser-Busch InBev	BE	2013-2020	since 09/2010	since 09/2009	-
ASML	NL	2013-2020	since 08/2016	since 06/2012	-
Assicurazioni Generali	IT	2013-2020	-	08/1998-09/2016	-
AstraZeneca	GB	2013-2020	since 09/2000	-	-
AXA	FR	2013-2020	since 09/2013	since 08/1998	-
Banco Santander	ES	2013-2020	09/1999-09/2020	since 09/1999	-
Barclays	GB	2013-2020	04/1999-12/2018	-	-
BASF	DE	2013-2020	since 09/2002	since 09/1999	-
Bayer	DE	2013-2020	since 09/2008	since 08/1998	-
BBVA	ES	2013-2020	08/1998-09/2019	08/1998-09/2020	-
BG Group	GB	2013-2015	04/2009-02/2016	-	Takeover by Royal Dutch Shell in 2016
BHP Group	AU/GB	2013-2020	09/2006-09/2015 & since 07/2020	-	Renamed from “BHP Billiton” to “BHP Group” in 2017/2018
BMW	DE	2013-2020	-	since 09/2010	-
BNP Paribas	FR	2013-2020	since 09/2000	since 09/1999	-
BP	GB	2013-2020	since 08/1998	-	-
British American Tobacco	GB	2013-2020	since 09/2008	-	-
BT Group	GB	2013-2020	09/2014-09/2017	-	-
Carrefour	FR	2013-2020	-	08/1998-09/2016	-
Credit Suisse	CH	2013-2020	06/2013-08/2016	-	-
CRH	IE	2013-2020	-	09/2009-09/2014 & since 09/2016	-
Daimler	DE	2013-2020	since 10/1998	since 11/1998	-
Danone	FR	2013-2020	-	since 09/2000	-

Continued on next page

Table E.16 – continued from previous page

Firm	Country	Sample period (FY)	STOXX Europe 50	EURO STOXX 50	Comments
Deutsche Bank	DE	2013-2020	08/1998-08/2016	08/1998-09/2018	-
Deutsche Börse	DE	2013-2020	-	since 09/2019	-
Deutsche Post	DE	2013-2020	-	since 09/2013	-
Deutsche Telekom	DE	2013-2020	since 08/1998	since 08/1998	-
Diageo	GB	2013-2020	since 08/1998	-	-
E.ON	DE	2013-2020	08/1998-09/2013	08/1998-09/2018	-
Enel	IT	2013-2020	since 09/2019	since 03/2000	-
ENGIE	FR	2013-2020	07/2008-03/2013	since 07/2008	Renamed from GDF Suez in 2015
Eni	IT	2013-2020	08/1998-07/2020	since 08/1998	-
EssilorLuxottica	FR	2013-2020	-	since 6/2012	Essilor merged in 2018 with Luxottica (IT), Renamed to EssilorLuxottica
Flutter Entertainment	IE	2016-2020	-	since 09/2020	Merged in 2016 from Betfair (GB) and Paddy Power (IE) to Paddy Power Betfair; Renamed to Flutter Entertainment in 2019
Fresenius	DE	2013-2020	-	09/2015-09/2020	-
GlaxoSmithKline	GB	2013-2020	since 08/1998	-	-
Glencore	JE	2013-2020	06/2013-09/2015 & 05/2017-09/2019	-	-
HSBC	GB	2013-2020	since 09/1999	-	-
Iberdrola	ES	2013-2020	since 09/2019	since 09/2003	-
Imperial Brands	GB	2013-2020	12/2011-09/2013 & 09/2015-09/2018	-	2016 Renaming of Imperial Tobacco to Imperial Brands
Inditex	ES	2014-2020	-	since 09/2011	No data available for FY 2013
ING Groep	NL	2013-2020	08/1998-05/2020	since 08/1998	-
Intesa Sanpaolo	IT	2013-2020	since 09/2015	since 01/2007	-
Kering	FR	2013-2020	since 05/2020	since 09/2018	-
KONE	FI	2013-2020	-	since 09/2020	-
Linde SE	IE	2018-2020	since 12/2018	since 10/2018	Merger in 10/2018 of “Praxair” (US) and “Linde AG” (DE)
Lloyds Banking	GB	2013-2020	09/2013-09/2020	-	-
L’Oréal	FR	2013-2020	since 02/2016	since 08/1998	-
LVMH	FR	2013-2020	since 09/2011	since 08/1998	-
Münchener Rück	DE	2013-2020	-	since 09/1999	-
National Grid	GB	2013-2020	since 09/2011	-	-
Nestlé	CH	2013-2020	-	since 08/1998	-

Continued on next page

Table E.16 – continued from previous page

Firm	Country	Sample period (FY)	STOXX Europe 50	EURO STOXX 50	Comments
Nokia	FI	2013-2020	-	08/1998-03/2013 & since 09/2014	-
Novartis	CH	2013-2020	since 08/1998	-	-
Novo Nordisk	DK	2013-2020	since 06/2015	-	-
Orange	FR	2013-2020	-	08/1998-09/2020	-
Pernod Ricard	FR	2013-2020	-	since 09/2020	-
Philips	NL	2013-2020	-	since 08/1998	-
Prosus	NL	2020	since 11/2020	since 09/2020	Founded in 2019; first annual report for FY 2020
Prudential	GB	2013-2020	since 09/2014	-	-
Reckitt Benckiser	GB	2013-2020	since 06/2012	-	-
RELX Group	GB	2013-2020	since 09/2019	-	-
Repsol	ES	2013-2020	-	08/1998-09/2015	-
Richemont	CH	2013-2020	12/2012-09/2016	-	-
Rio Tinto	AU/GB	2013-2020	since 09/2005	-	-
Roche	CH	2013-2020	since 09/1999	-	-
Royal Dutch Shell	GB	2013-2020	since 12/2005	-	-
RWE	DE	2013-2020	-	08/1998-09/2015	-
Safran	FR	2013-2020	since 09/2018	since 09/2015	-
Saint-Gobain	FR	2013-2020	-	09/2001-09/2018	-
Sanofi	FR	2013-2020	since 09/2007	since 09/1999	-
SAP	DE	2013-2020	since 09/2003	since 07/2004	-
Schneider Electric	FR	2013-2020	since 03/2013	since 09/2007	-
Siemens	DE	2013-2020	since 08/1998	since 08/1998	-
Société Générale	FR	2013-2020	-	08/1998-09/2020 -	
Standard Chartered	GB	2013-2020	09/2010-06/2015	-	-
Syngenta	CH	2013-2016	09/2016-05/2017	-	Takeover by ChemChina in 2017
Telefónica	ES	2013-2020	08/1998-09/2019	08/1998-09/2020	-
Total	FR	2013-2020	since 11/1999	since 06/1999	-
UBS	CH	2013-2020	since 11/2014	-	-
Unibail-Rodamco-Westfield	FR	2018-2020	-	02/2010-09/2019	Takeover of “Westfield” (AUS) in 06/2018
UniCredit	IT	2013-2020	-	08/1998-09/2016	-
Unilever Group	GB	2013-2020	since 09/2011	-	-
Vinci	FR	2013-2020	since 08/2016	since 09/2007	-
Vivendi	FR	2013-2020	-	since 08/1998	-
Vodafone	GB	2013-2020	since 09/1999	-	-
Volkswagen	DE	2013-2020	-	since 09/2011	-

Continued on next page

Table E.16 – continued from previous page

Firm	Country	Sample period (FY)	STOXX Europe 50	EURO STOXX 50	Comments
Vonovia	DE	2015-2020	-	since 09/2020	Merger in 08/2015 of “Deutsche Annington” (DE) and “GAGFAH” (LU)
Zurich Insurance	CH	2013-2020	since 09/2010	-	-

Notes: Index time specification includes changes up until FY2020