Econometric analysis of income distribution and wage differentials in the German labour market

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

vorgelegt von

MIRIAM CAROLIN SUSANNE STURM
(geb. Morlock)

Tübingen, 2025

Tag der mündlichen Prüfung:	25.02.2026
Dekanin:	Prof. Dr. Taiga Brahm
Dekan:	Prof. Dr. Dominik Papies
1. Gutachter:	Prof. Dr. Martin Biewen
2. Gutachter:	Prof. Dr. Joachim Grammig

Publication notes

Chapter 2 is based on Biewen, Martin and Miriam Sturm (2022): Why a labour market boom does not necessarily bring down inequality – Putting together Germany's inequality puzzle. *Fiscal Studies*, 43(2), 121-149.

Chapter 3 is based on Biewen, Martin, Pascal Erhardt, Bernd Fitzenberger and Miriam Sturm (2025): Selectivity-corrected wage distributions and the evolution of the German gender wage gap. *Unpublished manuscript, University of Tübingen.*

Chapter 4 is based on Sturm, Miriam (2025): Secular changes in educational attainment and the quality of the highly skilled – Evidence from Germany. *Unpublished manuscript, University of Tübingen.*

Funding

The research for this doctoral thesis was in parts funded through DFG priority programme 1764 "The German Labour Market in a Globalised World: Challenges through Trade, Technology and Demographics" and DFG projects BI767/3-1 and BI767/3-2. Chapter 2 is part of a series of research projects contributing to the 6th Report on Poverty and Wealth by the German Federal Ministry of Labour and Social Affairs. Financial support by the DFG as well as by the German Federal Ministry of Labour and Social Affairs is gratefully acknowledged.

Acknowledgements

Anyone who has read this dissertation will have noticed that it asks many "what if" questions. Yet these only constitute a tiny fraction of all the "what ifs" I've asked myself over this long educational journey of mine. Perhaps I got into econometrics because of my urge to wonder about potential outcomes, always wanting to know what could have been or what could happen in any conceivable parallel universe. Over the last few years, many of these "what ifs" have revolved around this dissertation, which has often stretched my husband's patience to its limits. Luckily for both of us, as of today my dissertation is one less thing for me to worry about. I am officially done, and I am immensely grateful to many people who have all helped me to get to this point.

First of all, I am deeply indebted to my supervisor and co-author Martin Biewen. Thank you, Martin, for being a great teacher from the first EDA lecture to the day I finished this dissertation; for your trust in my abilities and your continuous encouragement; and not least for making my work at the chair compatible with our at times challenging Kita opening hours. I am beyond thankful for all your support over these past years. I am also very grateful to my second supervisor Joachim Grammig for his dedication to teaching, which I greatly benefitted from as a student; for his feedback and encouragement during my time as a PhD; and for regularly lifting all our spirits with his contagious enthusiasm about econometrics.

Over the years I got to know a bunch of wonderful people with whom I shared the ups and downs of academia, especially the members of the two Econometrics chairs. I will particularly remember our "exile" below the dentist's office, which forced us to move closer together and in the end resulted in new friendships. Thank you for all the shared lunches and coffees; for the professional and private exchange; and for never letting me feel alone with my "what ifs". Special mention goes to my office roomy Jakob as well as to Marian and Pascal for their expertise as co-authors and all the laughs along the way; to our Hiwis for their invaluable work; and to Lea, Alexandra and Sylvia for being the heart of our two chairs that always kept things running.

I would further like to thank Tobias Brändle and Bernhard Boockmann for sparking my interest in policy evaluation and data analysis during my internship at the IAW, which was when a PhD in statistics became a potential outcome I could see for myself in the first place. Thank you also to Rolf Kleimann for the close and trusting collaboration for the 6th Report on Poverty and Wealth by the Federal Ministry of Labour and Social Affairs; and to Bernd Fitzenberger for everything I learned while working on our joint paper. It was a privilege to have been part of both of these projects.

There are also countless people outside of my little research bubble who have contributed to this thesis in various ways, and to whom I am deeply grateful. To the wonderful people working at my children's daycare and the other parents in our group who have helped each other out over these past years; to my friends who thankfully at some point cut their jokes about how I was an "eternal student" and "not really working"; and not least to my parents who gave me an infinite range of possibilities without pushing me into any direction, and who – along with my father-in-law – have been an immeasurable help as babysitters especially since our Kita's opening hours were cut. Turning the potential outcome of me pursuing – let alone actually getting – a PhD into an observed one would never have been possible without them, and I am infinitely grateful for everything they've done for me.

But most of all this is for Pascal, who was by my side in the very first lectures we attended as first-year students and has been ever since; and for my little whirlwinds Emmi and Luki, the only humans on this earth who always, always manage to get me back on my feet just by being their cheerful, goofy selves. The three of you are the reason for which I've come to realise that getting hung up on potential outcomes can also be terribly unproductive, even kind of pointless – at least when it comes to real life. Out of all my potential outcomes, the one life had in store for me has the three of you in it, and that's all that really matters. As long as I have you, I actually – just this once – don't care about what could have been.

Miriam Sturm
Tübingen, November 2025

Contents

List of figures

List of tables

CHAPTER 1

Dissertation introduction

Questions of distribution and differences in outcomes for individuals are not the sole part of economics – to suggest that would be unwarranted – but they are an essential part.

Anthony Atkinson, *Inequality: What can be done?*

Whether in politics, in academia or at the dinner tables of families all over the world, there is a widespread consensus that inequality is one of the great challenges of our time. Inequality is a rather broad term per se: While some may primarily worry about unequal chances to succeed in the education system, others may think of inequality in political participation. Inequality can also be mirrored in disparate access to justice or health care, or simply in the way some individuals are treated differently than others in everyday life. Yet the first association with it – especially in political debates – is typically that of *income* inequality. Surely, part of this is because economic variables such as wages or asset returns can be comparatively easily pinned down in numbers while other types of inequalities are somewhat less palpable. Another crucial point, though, is the empirical observation that income inequality is deeply interwoven with other inequalities, thus spilling over into other societal domains beyond economics.

It is therefore not surprising that economic research on distributional matters has prospered over the course of the last decades, which have been characterised by growing inequality levels in many industrialised countries. In the US, for instance, the effort to understand the sources of wage inequality was significantly enhanced after its surge in the late 1970s, fuelled unquestionably by the growing availability of micro-level data which offered a whole new range of possibilities at the time. However, another key driver behind the soaring research interest in unequal pay was the fact that the rising inequality levels became increasingly tangible in individuals' lives, not least since "widening wage structure has meant widening family income and consumption inequality and associated social problems" (Katz and Autor, 1999, p. 2). Empirical evidence indeed suggests that income inequality is related – inter alia – to higher crime rates (Dahlberg and Gustavsson, 2008; Freeman, 1999; Kelly, 2000), unequal or detrimental health outcomes (Chetty et al., 2016; Pickett and Wilkinson, 2015), and lower educational mobility (Blanden, 2013; Corak, 2013; Jerrim and Macmillan, 2015). Furthermore, countries in which (perceived) inequality levels are high tend to be characterised by lower degrees of social cohesion (Alesina and La Ferrara, 2000; Gould and Hijzen, 2016) and trust in meritocratic processes and fairness (Kuhn, 2019).

Arguably, the perception of what is distributionally fair hinges on a variety of aspects that are often – in one way or another – connected to normative values. Even from a neutral perspective, one can reach vastly different conclusions on the existing amount of inequality in a society depending on various factors, such as the inequality index used, the income-receiving unit defined, and not least on the precise measurement of income (Cowell, 2000; Jenkins, 2024). This may range from high-frequency (e.g., hourly) wages to after-government-after-tax disposable (household) income. Economists generally agree that which components to include into an assessment of inequality should first and foremost be governed by the policy question at hand, although the choice also needs to be made based on data availability in practice (Armour et al., 2013). This is also why many empirical contributions on inequality focus on income in the first place instead of also considering wealth, which may crucially influence individual well-being, but on which there is typically a lack of transparent and reliable data (Zucman, 2019).[1]

[1] For a recent historical review of the German wealth distribution that is not solely based on survey data – which is prone to misreporting particularly in the case of wealth –, see Albers et al. (2025).

This dissertation aims to contribute to the empirical literature on income inequality in Germany, moving from a particularly broad definition to more specific measures of between-group inequalities: Chapter 2 is concerned with inequality in disposable income and its various ingredients, which includes – to name just a few of them – labour income, capital income, as well as private and government transfers. This income definition also considers the given household structure, accounting for the fact that the financial resources of a household are typically shared among its members. While this measure is tightly connected to personal consumption possibilities and well-being, it is not particularly well-suited for assessing inequalities between certain groups of individuals (e.g., between genders). In contrast, chapters 3 and 4 are solely dedicated to individual labour income, which – in the German case – constitutes the lion's share of total income across large parts of distribution (Drechsel-Grau et al., 2022).[2] More specifically, these two chapters use administrative daily wage data to look more closely at wage differentials between genders (see chapter 3) and individuals with different levels of educational attainment (chapter 4), both of which play a crucial role in shaping the distribution of wages and, hence, the final distribution of incomes.

While each of the three chapters looks at inequality from a different angle, they have in common that they are all based on German data, making it indispensible to outline the general distributional trends in Germany over the last decades. Fuchs-Schündeln et al. (2010) describe the evolution of inequality in Germany until the mid 2000s as a "tale of two countries". The first of these two is the country that existed prior to the German reunification, which saw remarkably stable inequality throughout the 1980s with only some increases in the upper tail of the distribution (Antonczyk et al., 2018; Dustmann et al., 2009). The second part of the tale, however, tells the story of a post-unification Germany witnessing rising and quickly accelerating wage and earnings inequality in all parts of the distribution that happened against the backdrop of slow average wage growth, declining incomes in the lower half of the distribution, and increasing incomes at the top (Drechsel-Grau et al., 2022; Dustmann et al., 2009; Fuchs-Schündeln et al., 2010). This trend occurred somewhat delayed compared to the US, where inequality had already begun to rise sharply two decades earlier and where its growth finally

[2] Drechsel-Grau et al. (2022) define total income as the sum of labour and non-labour income. The latter comprises income from self-employment, rental income and business income, which is the dominant source of income in the right tail.

slowed down towards the late 1990s (Autor et al., 2008; Katz and Autor, 1999) while it started to speed up around that same time in Germany (Card et al., 2013; Dustmann et al., 2009). Fuchs-Schündeln et al. (2010) document the surge in inequality prior to the millennium across a variety of income measures, although it was less pronounced for post-government measures accounting for the redistributive effects of the German tax and transfer system.

Inequality also climbed steadily after the millennium, and it continued to do so until the mid to late 2000s. These years then mark a turning point, with wage inequality having stagnated or even slightly declined from roughly 2009/10 onwards (see, e.g., Felbermayr et al., 2016; Möller, 2016) whereas inequality in disposable (equivalised) incomes had already remained rather constant since 2005/06 (Biewen et al., 2019). Drechsel-Grau et al. (2022) observe a similar trend reversal when considering annual labour incomes, which they find to have become more dispersed leading up to the Great Recession, while incomes grew throughout the distribution – and most notably for women in the bottom half – afterwards. This consequently led to a decline in inequality across the whole distribution, with the exception of the extreme right tail where earnings inequality has kept rising to the day (Drechsel-Grau et al., 2022).

Unsurprisingly, the periods of rising and stagnating inequality were characterised by markedly different economic conditions in general. In the 1990s and early 2000s, Germany struggled with the repercussions of the economic shock of the reunification, yielding the country its unpopular nickname as the so-called “sick man of Europe” (Dustmann et al., 2014). The record unemployment levels in particular were of great public concern and finally led to a series of profound labour market reforms – also referred to as the “Hartz reforms” – targeting activating labour market policies and a flexibilisation of the German labour market. Shortly after the latter went into effect (i.e., in the mid 2000s), the formerly mentioned “sick man” underwent a transformation to an “economic superstar” (Dustmann et al., 2014) whose labour market kept chasing its own records until the COVID-19 pandemic: Unemployment fell from 12% in 2004 to below 6% in 2018 and labour force participation rose significantly, though in a much more pronounced way for women where it increased from 55 to over 70% (Drechsel-Grau et al., 2022).

In spite of the temporal proximity of the "Hartz reforms" and Germany's economic upsurge, it remains controversial to the day as to whether these have been the primary driver of what Burda (2016) refers to as the German "labour market miracle".[3] Irrespective of what caused it in the first place, the sheer magnitude of Germany's employment boom may have raised the expectation that the country would finally see a perceptible decline in disposable income inequality. However, as previously mentioned, Germany in fact only witnessed a very modest reduction in inequality – if at all – between the mid 2000s and the onset of the pandemic. This somewhat puzzling observation is at the core of chapter 2 of this dissertation, which closely investigates the question whether the observed employment gains contributed to a decline in inequality in net equivalised household incomes at all and – if so – which other factors have outweighed and eventually masked this equalising effect.

Shifting the focus from the household to the individual level, Germany's employment boom has taken on somewhat different forms for men and women, respectively. At the extensive margin, female labour market participation increased substantially more than it was the case for men (see above). A look at the intensive margin further reveals that the rising employment shares of women did not translate into a corresponding increase in the female full-time share, but were predominantly driven by women entering part-time employment out of non-participation: The female part-time share (relative to the entire female prime-aged population) increased from about 20% in 2005 to over 30% within only one decade, whereas its male counterpart rose from a modest 2.5% to just over 5% over the same time span (Carrillo-Tudela et al., 2021). Consequently, the part-time share among *employed* women has remained at its traditionally high level in international comparison, with roughly every second woman – compared to only one in eight men – working part-time in 2023 (Statistisches Bundesamt, 2024).

The discrepancy between male and female labour force participation – as well as the intensity thereof – is no purely German phenomenon, and has posed conceptual chal-

[3] A number of studies attest an important role to the "Hartz reforms" in explaining Germany's employment boom (Bradley and Kügler, 2019; Felbermayr et al., 2016; Hartung et al., 2018; Launov and Wälde, 2013). Others are somewhat more sceptical in attaching a causal meaning to them (Dustmann et al., 2014; Hoffmann and Lemieux, 2016).

lenges for applied labour market research all over the globe.[4] In fact, the comparatively low (full-time) participation rates of women are the main reason why much of the literature exploring the determinants of wages focusses exclusively on the wage schedules of men, arguing that wage equations estimated based on a specific, potentially non-representative fraction of the entire population may involve significant (self-)selection bias in the case of women (see the seminal work of Heckman, 1974, 1979). This same notion has also called the validity of conventional measures of the gender pay gap[5] into question since these often ignore the potential for selection bias, eventually leading to a distorted perception of the extent of unequal pay between men and women.[6]

Over the years, increasingly sophisticated techniques have emerged to recover not only averages, but wage quantiles at different parts of the distribution (Albrecht et al., 2009; Biewen et al., 2020; Chzhen and Mumford, 2011) or even whole wage distributions (Arellano and Bonhomme, 2017; Chen et al., 2024) that are purged of selection. In earlier contributions, the application of methods that account for selectivity was almost exclusively targeted at women, while male selection was long deemed negligible. This view has somewhat shifted in recent years against the background of secular trends on the labour market. Blau et al. (2024), for instance, highlight the observation that the decline in the US gender wage gap coincided with a convergence of male and female labour force participation rates. The authors trace the simultaneity of these trends back to shifting selection patterns of both genders caused by their changing participation. Indeed, a growing number of contributions have documented sizeable female *and* male selection bias (especially around the years of the Great Recession, see Dolado et al., 2020; Elass, 2024), though no comparable evidence exists for Germany so far.

From an econometric perspective, a remaining shortcoming in the literature lies in its restrictiveness regarding estimation of the parameter determining the direction and strength of the selection bias, which is typically assumed to be constant across the

[4]The evolution of the role of women on the labour market is documented in detail by the pioneering research of Claudia Goldin (see, e.g., Goldin 2006, 2014, 2021) for which she received the Nobel Prize in Economics in 2023.

[5]For recent reviews of the vast literature on the gender pay gap, see Kunze (2018) and Olivetti et al. (2024).

[6]Whether uncorrected measures over- or underestimate the "true" (selection-corrected) gender pay gap is not unambiguous in the literature: Numerous contributions report a growing gap once accounting for selection (see, e.g., Albrecht et al., 2009; Biewen et al., 2020; Picchio and Mussida, 2011) while others obtain the opposite result (Chernozhukov et al., 2025; Elass, 2024).

entire wage distribution. It was only recently that Chernozhukov et al. (2025) proposed a distribution regression approach allowing explicitly for different signs and degrees of selection on unobservables at different parts of the distribution. In an assessment of the German gender pay gap, chapter 3 of this thesis exploits this extraordinarily flexible framework in order to correct the male and female full- and part-time wage distributions for selection, unveiling considerable heterogeneities from the bottom to the top.

Methodically, large parts of this dissertation draw from the rich decomposition literature in labour economics (see Fortin et al., 2011, for a comprehensive summary). Decomposition methods constitute helpful tools when the goal is to identify the drivers of observed trends in inequality (as in chapter 2) or the sources of wage differentials such as the gender pay gap (see chapter 3). Building on the work of Oaxaca (1973) and Blinder (1973), numerous techniques have been developed that break down overall differences in wages Y into an effect resulting from differences in the marginal distribution of covariates X (composition effect) on the one hand, and one resulting from differences in the returns β received for these characteristics (wage structure effect) on the other.

A related question in the inequality literature is to what extent observed changes in selected inequality indices are the result of changes in inequality within as opposed to between groups. This is best illustrated based on the unconditional variance of Y, which – by its appealing statistical property of decomposability – can alternatively be written as the sum of a within-group and a between-group component:

$$\mathrm{Var}(Y) = \mathrm{E}[\mathrm{Var}(Y|X)] + \mathrm{Var}[\mathrm{E}(Y|X)].$$

The first component thereby measures the wage dispersion among individuals who are comparable in terms of their characteristics X, and is therefore also referred to as the residual variance. In the US and Germany alike, residual wage dispersion has been shown to closely mirror the aforementioned trends in overall wage dispersion (Autor et al., 2008; Fuchs-Schündeln et al., 2010). Another interesting finding in the literature is that residual inequality tends to be higher at the top compared to the bottom of the distribution (see Brüll and Gathmann, 2020, for the German case), a phenomenon that is often attributed to a larger heterogeneity among those individuals with the highest levels of educational attainment (Lemieux, 2006b; Martins and Pereira, 2004).

This latter observation is particularly important against the background of the striking educational expansion of the workforce witnessed in many industrialised countries over the last decades. Germany constitutes no exception from this trend: By way of example, the university share among 35- to 44-year-old men has more than doubled from roughly 10% to over 25% between 1975 and 2018 (Statistisches Bundesamt 1976, 2019). Likewise, the first-year student ratio has been steadily rising, thus contributing to a continuation of the observed "academisation" of the workforce (Wolter and Kerst, 2015). While there is a broad consensus in the literature that comparable trends in educational attainment are important in explaining the evolution of residual inequality, the underlying mechanisms are somewhat more controversial. In case of the US, for instance, Lemieux (2006a) argues that much of the pronounced increase in residual wage dispersion in the 1980s was in fact driven by a heteroscedasticity effect linked to educational expansion: If the share of highly educated individuals is rising – as it had been in the US – and unobservable skill or "ability" is generally more dispersed among this group of workers, this will equally cause residual inequality to rise. This view was later challenged by Autor et al. (2008), who argue that this merely proportional effect explains a relatively small part of the observed trends in residual inequality compared to price effects.

Apart from these alternative explanations, other researchers claim that educational expansion may also have an immediate impact on the distribution of unobserved skill within educational groups, most notably among those with the highest level of educational attainment. This hypothesis is based on the following argument: Assuming that the distribution of ability in the general population is approximately fixed over time, rising college shares may draw individuals located towards increasingly lower percentiles of the ability distribution into higher education, leading to a rising dispersion of ability and – consequently – also of wages within this group of workers (Carneiro and Lee, 2011; Juhn et al., 2005). At the same time, this exact mechanism may – ceteris paribus – result in a lower *between-group* inequality (see second component of the variance formula) caused by a somewhat less favourably selected group of college graduates. This component thereby captures average wage differences between worker groups with characteristics X, i.e., average skill premia.

Indeed, a number of contributions in the vast literature on the returns to education in Germany document significant cohort patterns in skill premia (Antonczyk et al., 2018; Boockmann and Steiner, 2006; Reinhold and Thomsen, 2017) and hypothesise about their link to potential ability shifts induced especially by the rising tertiary education shares. Given the importance of educational upgrading in shaping the German wage distribution (see, e.g., Biewen and Seckler, 2019), chapter 4 of this dissertation explores this link in more detail.

To conclude this introduction, the following paragraphs provide an overview of the three studies this doctoral thesis is based upon.

Why a labour market boom does not necessarily bring down inequality –
Putting together Germany's inequality puzzle.

Chapter 2 of this dissertation provides a comprehensive and holistic decomposition analysis of the factors that have shaped the distribution of net equivalised incomes in Germany between 2005/06 and 2015/16, focussing in particular on the aforementioned "labour market miracle" that unfolded over these years. Motivated by the notion that not even the ever-increasing employment levels have led to a substantial decline in disposable income inequality, we investigate to what extent the observed gains in full-time, part-time and marginal part-time employment have had an effect on the distribution of incomes at all. We thereby contribute to a comparatively small body of literature dealing with disposable income, which has thus far struggled to establish a clear link between the evolution of inequality on the one hand and the developments on the labour market on the other (Biewen et al., 2019; Peichl et al., 2018). Our results imply that while the flourishing labour market conditions alone did have a significant equalising effect, the latter was considerably dampened by the tax and transfer system and additionally masked by the disequalising impact of other factors such as immigration (resulting in an inflow of low-income households) and changes in household characteristics (most notably population ageing and educational expansion).

Selectivity-corrected wage distributions and the evolution of the German gender wage gap

Against the background of the rising employment levels that are at the core of the previous analysis, chapter 3 investigates whether these have been accompanied by a changing composition of the German workforce regarding unobservable characteristics. More specifically, our goal is to examine how male and female selection patterns into full- and part-time work have changed over time, and how these changes have affected the evolution of gender wage gaps in both employment forms between 2000-2005 (recession) and 2012-2017 (boom). One of the main contributions of our study is that it is the first to employ the approach recently proposed by Chernozhukov et al. (2025), which constitutes a methodological breakthrough since it explicitly allows for different signs and degrees of selectivity across the distribution. Our results indeed reveal sizeable heterogeneities: Full-time men tend to be positively selected at the bottom (potentially due to the generous social safety net) and negatively in the rest of the distribution – a trend that intensified over the course of the employment boom. In contrast, full-time women are generally negatively selected, and increasingly so towards the top. This may be driven by assortative matching and the German tax and transfer system that discourages secondary earnings. Part-time men also turn out to be negatively selected, while female sorting into part-time exhibits a more complex structure, switching from negative at the bottom to positive in the higher percentiles. Full- and part-time gender wage gaps have narrowed due to declining differences in unobserved selectivity between men and women, but also as a result of improved observables of full-time women and converging wage returns in male and female part-time work.

Secular changes in educational attainment and the quality of the highly skilled – Evidence from Germany

Besides the flourishing labour market participation rates, one of the most salient developments on Germany's labour market has been the ongoing educational upgrading of its workforce: Over the last decades, the tertiary education share has risen continuously, while dual vocational training has become less popular for successive cohorts. In light of a growing body of literature that simultaneously documents inter-cohort differences in educational attainment *and* in educational returns, the question emerges as to whether the rising college shares have drawn individuals with increasingly lower levels of unobserved skill or "ability" into tertiary education. Although a number of studies have hypothesised about the existence of such effects, they have struggled to establish a clear link between educational expansion and potential ability shifts since changing educational shares necessarily translate into changing relative supplies, making it hard to isolate ability effects from supply effects. Similar to Carneiro and Lee (2011) for the US, chapter 4 aims to resolve this issue by exploiting regional variation, comparing the wages of college educated individuals who face the same labour market conditions, but who received their college entry certificate ("Abitur") in different federal states. The results provide tentative evidence that college educated persons from regions with relatively high "Abitur" shares tend to receive lower wages than their counterparts from regions with lower shares, potentially pointing to a declining average ability as relatively more persons formally qualify for tertiary education. The analysis further indicates potential spillover effects on those holding a vocational training degree. However, these are not as pronounced as for the college educated, implying that the college premium would have been higher in absence of the observed educational expansion.

CHAPTER 2

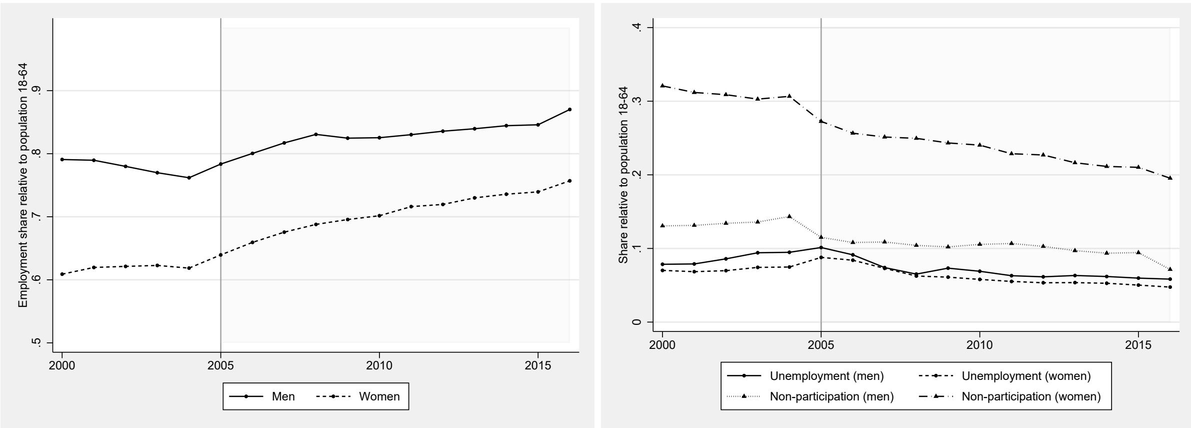
Why a labour market boom does not necessarily bring down inequality – Putting together Germany's inequality puzzle*

2.1 Introduction

Following reunification in 1990, Germany had to face difficult economic conditions throughout the 1990s and the early 2000s: low economic growth, a high fiscal deficit and increasing unemployment. In the mid 2000s, however, the so-called 'sick man of Europe' took off and experienced an unprecedented employment boom that has been

*This chapter is based on: Biewen, Martin and Miriam Sturm (2022): Why a labour market boom does not necessarily bring down inequality – Putting together Germany's inequality puzzle. *Fiscal Studies*, 43(2), 121-149. We thank the editor, James Ziliak, three anonymous reviewers, as well as Christina Gathmann, Regina Riphahn, Judith Niehues and participants at the Annual Conference of the European Economic Association EEA 2020, the Annual Meeting of the Verein für Socialpolitik 2020 and the Annual Meeting of the European Society for Population Economics ESPE 2021 for valuable comments and suggestions.

chasing its own records in recent years (Dustmann et al., 2014). Not even challenging events such as the global financial crisis in 2008–09 or the drastically increased immigration since 2014 (often referred to as the 'refugee crisis') have interrupted Germany's economic upsurge, which was only stopped by the global COVID-19 crisis starting in 2020.



(a) Employment shares **(b)** Unemployment and non-participation shares

Figure 2.1 – Aggregate employment, unemployment and non-participation shares by gender, 2000–16
(Source: Federal Statistical Office)

The magnitude of this boom is shown in figure 2.1. After several years of stagnation, employment rates began to rise significantly for men, and even more so for women. The boom drastically reduced unemployment and boosted labour market participation, particularly female participation. A number of previous contributions have examined the structure of these employment gains. For instance, Rothe and Wälde (2017) claim that a large number of the unemployed who found a job during the boom did not go into full-time work. Rather, they observe a substantial increase in part-time employment and non-standard work (e.g., marginal employment). However, Ehrich et al. (2018) and Carrillo-Tudela et al. (2021) emphasise that the boom increased participation in general, drawing individuals into the labour market who would not have participated otherwise. This was particularly true of women who often entered part-time or marginal employment from non-participation.

At the same time, and as shown below, income inequality in Germany first stagnated after the onset of the boom in 2005 but then followed a slight upward trend from 2010 onwards. Given the nature of the boom, this constitutes somewhat of a puzzle. In view of the drastic reduction in unemployment from over five million individuals to around half this value, and the additional participation in part-time and marginal employment, the boom should have massively benefitted those at the bottom of the income distribution, leading to a reduction of income inequality.

A small number of previous contributions have considered the development of income inequality in Germany after 2005. Peichl et al. (2018) and Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2019) document the evolution of inequality measures for disposable incomes over the same period as we do in this paper. These studies provide evidence for first stagnating and then slightly increasing inequality but do not attempt to relate this finding to other changes such as the employment boom. Biewen et al. (2019) consider the period 2005–10 but struggle to establish an effect of both the massive expansion and the compositional changes in employment after 2005 on the resulting distribution of household disposable incomes. Dustmann et al. (2022) present an analysis of inequality trends based on the alternative dataset *Einkommens- und Verbrauchsstichprobe (EVS)*. They also find slightly rising inequality between 2003 and 2013 but their focus is on the role of housing expenditures on inequality.

In this paper, we aim to make the following contributions. First, we provide more evidence on the exact structure of the 'German labour market miracle' (Burda, 2016), which has drawn a lot of attention in the literature.[1] We show that the boom was not a shift from full-time to part-time employment but involved net gains in both categories with the strongest component coming from the expansion of female part-time employment. Second, we present an explicit analysis of the effects of the boom on the distribution of net incomes based on rich microdata. Such an analysis is challenging because such effects depend on who exactly gained from the boom, how employment structures changed within households, and how the tax and transfer system transformed gains into net incomes. We explicitly consider heterogeneous effects of the boom on households by modelling in detail changes in labour market participation conditional on a rich set of individual and household characteristics. As employment trends are not the only source

[1] A number of alternative explanations for the boom have been proposed: wage restraint and de-unionisation (Dustmann et al., 2014; Kügler et al., 2018), export boom (Dauth et al., 2017; Dauth et al., 2021), Hartz reforms (controversial, see Akyol et al., 2013; Bradley and Kügler, 2019; Burda and Seele, 2020; Hartung et al., 2018; Hochmuth et al., 2021; Hutter et al., 2022; Launov and Wälde, 2013; Launov and Wälde, 2016), changes in early retirement (Riphahn and Schrader, 2020) and in parental leave legislation (Bick, 2015; Geyer et al., 2015).

of changes in the distribution of net incomes, we also consider the effects of confounding factors such as changes in pay structures, changes in the composition of the population, changes due to immigration, changes in other income sources such as capital income, and changes in the tax and transfer system.

Third, we contribute to a surprisingly small body of literature that analyses possible causes for changes in the distribution of *net incomes* – which is the income distribution relevant for welfare analysis and policy – but which is the complex result of a large number of elements such as employment, pay structures, household arrangements and institutional circumstances (Biewen and Juhasz, 2012; Blundell et al., 2018; Daly and Valletta, 2006; Hyslop and Maré, 2005; Jessen, 2019; Sologon et al., 2019). Much of the literature deals with *gross incomes* (often derived from tax records, e.g., Armour et al., 2013) or individual income components such as wages, which makes it difficult to assess the consequences for the final distribution of *net incomes* (see Armour et al., 2013, for a related point). An important takeaway from our analysis is the distinction between developments in pre- and post-tax/transfer incomes.

We reach the following conclusions. Despite the continuing trend of rising income inequality after 2005, the employment boom did have an equalising effect. This effect, however, was substantially dampened by the generous social security system, in particular unemployment insurance. One of the main purposes of this system is to insure income losses due to job loss or other unforeseen causes. On the positive side, this substantially alleviates the effect of economic downturns. On the negative side, however, this also reduces the effects of economic upturns on net incomes. The impact of the German social security system appears particularly strong in this respect. Our results suggest that even if the economic consequences of the COVID-19 shock were to reverse all the employment gains that occurred during the boom, this would only have modest effects on the distribution of net incomes. We further show that much of the employment boom took the form of additional part-time and marginal part-time work for women, both in single and in non-single households. Thus, our results demonstrate that a long expansion of employment may have a favourable effect on the distribution of incomes, even if much of the employment gains take the form of part-time and marginal part-time work, which are often seen as inferior forms of employment. Finally, we demonstrate that distributional effects of employment changes may be masked by other developments, making it hard to determine their exact magnitude. In our case, we show that the equalising impact of the boom was partly offset by immigration of individuals with low disposable incomes and by long-term compositional changes in the population (educational upgrading and population ageing).

The remainder of this chapter is structured as follows. In section 2.2, we describe the data underlying our study. Section 2.3 provides an overview of recent trends in the German income distribution. In section 2.4, we present and discuss our empirical results. We conclude in section 2.5. Appendix section A.3 contains a more detailed outline of our methods whose description in the main text is kept brief.

2.2 Data

Our study is based on the German Socio-Economic Panel (SOEP), a representative study of households living in Germany collected and maintained by the German Institute for Economic Research (DIW), see Goebel et al. (2019). In spite of the general limitations of survey data, the SOEP constitutes the only data source containing sufficient information for a study covering all relevant aspects of the distribution of net household incomes such as different income components, employment outcomes and socio-economic characteristics of all household members. Besides the SOEP core survey, we exploit the information in the SOEP migration samples as well as in the IAB-SOEP refugee sample to assess potential effects of immigration (see details below).

The focal point of our analysis is the distribution of annual net equivalised incomes between the years 2005/06 (when the employment boom set in) and 2015/16 (the most recent survey years with available income information at the time our study was carried out).[2] Our measure of net equivalised income is based on annual household net income

$$y = y_{Market} + y_{Pens} + y_{Trans} - ssc(y_{Labour}, y_{Pens}) - tax(y_{Tax}), \qquad (2.1)$$

where y_{Market} denotes the sum of all household members' annual market incomes (labour

[2] We pool years in order to increase statistical precision and to make our analysis less dependent on individual years as in Hyslop and Maré (2005) or Blundell et al. (2007).

income and capital incomes such as income from interest, dividends, rents[3]), y_{Pens} is the sum of all pension incomes (private and public), and y_{Trans} is the sum of public transfers received. Household public transfers include the full range of government transfers such as unemployment benefits, child benefits, student grants and subsistence allowances (among others; for simplicity we also include under this label transfers *between* private households). The terms $ssc(y_{Labour}, y_{Pens})$ and $tax(y_{Tax})$ represent deductions of social security contributions (pensions, health, unemployment and old age care insurance) as well as income taxes paid by the household. We compute both of these components for each household using our own income tax and social security contributions module, which is described in appendix section A.5. In order to focus only on real income changes, we inflate nominal income measures to prices of our most recent year 2016 (in the case of taxes and social security contributions we do this after the respective calculations). Finally, we equivalise annual net household income using the commonly used modified OECD equivalence scale, and we attribute the resulting equivalised income measure to each household member.

A big strength of a survey dataset such as the SOEP is the availability of individual income components, mostly at the individual level, see Grabka (2017). This is crucial for our purpose as we aim to counterfactually alter individual components such as labour incomes in order to determine their effect on the resulting distribution of annual household net incomes. An important ingredient to this analysis is the availability of summary calendar information on monthly employment activities in different categories (full-time, part-time, marginal part-time, unemployment) as well as information on income earned in different employment activities (main job, side job, self-employment).[4] The information on employment in full-time, part-time and marginal part-time work is based on the self-reports of the survey participants intending to provide a summary picture of all their employment activities during a given year. In particular, they include the possibility of cumulative, parallel and/or multiple employment spells. The distinction between full-time, part-time and marginal part-time work is made by the survey participants themselves, but we expect them to follow the fact that most full-time jobs

[3]Following common practice, we also include imputed rental values for owner-occupied housing and imputed social security contributions for civil servants in household market income.

[4]Given our interest in the distribution of annual incomes, we do not focus on hours worked or hourly wages apart from our distinction into full-time, part-time and marginal part-time work. See Carrillo-Tudela et al. (2021) for some information on changes in hours worked during the period under consideration.

in Germany have a contractual working week of 35 to 41 hours. Part-time jobs usually have much lower working hours. Marginal part-time work is typically either occasional or additional employment with few or irregular working hours or takes a standardised form with certain exemptions from taxes and social security contributions ('minijobs', typically 400 to 450 euros per month).

Based on the information in the monthly income and employment calendars, we construct for each individual the annual number of months worked in different employment categories (full-time, part-time, marginal part-time) along with the average monthly wage received in the respective category.[5] We include in our definition of employment both dependent and self-employment. Our construction is such that multiplying and adding up individuals' months worked and monthly wages yields the annual labour income of each individual as reported in the SOEP and allows us to separately change employment quantities and wage rates in our counterfactual simulations.[6]

Our analysis makes use of a large number of further characteristics at the individual and at the household level. In general, we distinguish between the following six different household types: (i) single pensioner households (65 years or older), (ii) multiple pensioner households (at least one household member 65 years or older and no household member under 55 years), (iii) single adults without children, (iv) multiple adults without children, (v) single adults with children, and (vi) multiple adults with children. Within households we consider detailed individual information on the household head and (if present) the partner or the second oldest adult in the household (gender, age, nationality, educational qualification in three categories, work experience in years, see table 2.2). For certain purposes, we also use information on individual employment histories (such as the number of months worked in different employment categories in the past three years, see below for more details). In addition to the characteristics of individual

[5]This requires some choices to reconcile the information in the employment and income calendars, see appendix A.2 for more details. Appendix A.2 also provides information regarding the comparability of SOEP employment data with those from other sources.

[6]The full use of employment information from the annual activity calendars of household members is an important difference to our previous study Biewen et al. (2019) which only used crude information on employment at the household level and only from the survey month (rather than over a full calendar year) along with descriptive information about different income measures over time. This turns out to be a crucial difference, as Biewen et al. (2019) failed to establish a clear relationship between employment changes and changes in the distribution of net incomes. Another important difference is that our earlier paper considered only the short time period 2005/06 to 2010/11, whereas the current paper covers the whole period of the economic upturn 2005/06 to 2015/16.

household members, we consider information on the number of children in the household in different age categories (0-3, 4-6, 7-17 years), the number of further adults in the household, and whether the household resides in East or in West Germany.

All our computations make full use of the SOEP sampling weights provided by the DIW, which ensure that our results represent the full German population. For statistical inference, we use bootstrapping, taking into account the repeated observation of the same households in different years and the clustering of individuals within households when computing bootstrap confidence intervals (Biewen, 2002).

2.3 General trends

Figure 2.2 displays inequality trends in equivalised net incomes since the year 2000. Consistent with previous contributions, the graphs show that income inequality first stagnated after the onset of the labour market boom in 2005 but then followed a slight upward trend from 2010 onwards. The upward trend after 2010 is present in the upper half of the distribution (percentile ratio P90/P50), but is even more pronounced in the lower half (percentile ratio P50/P10).

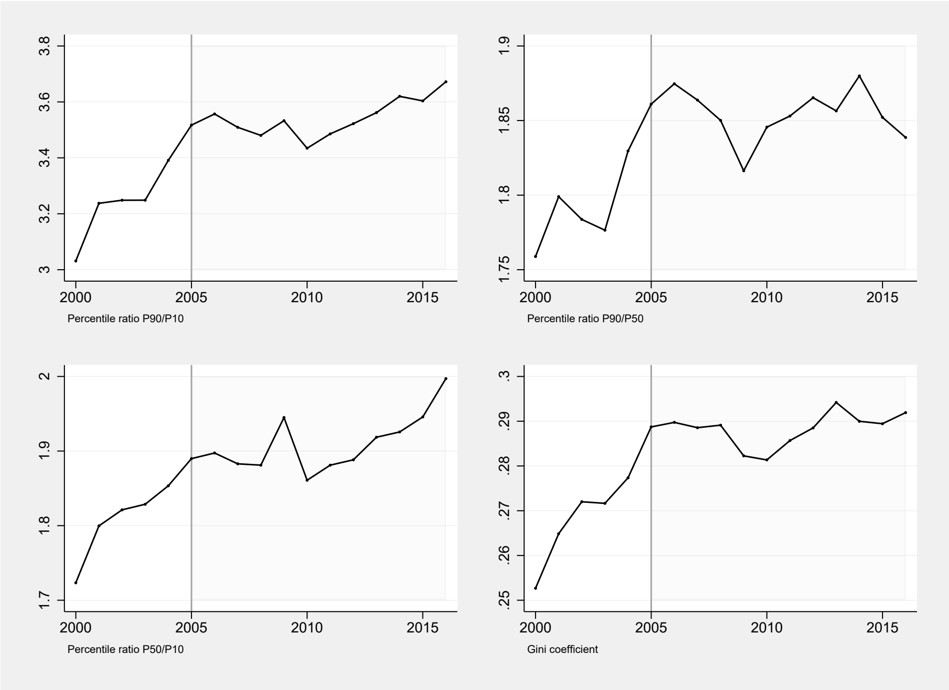
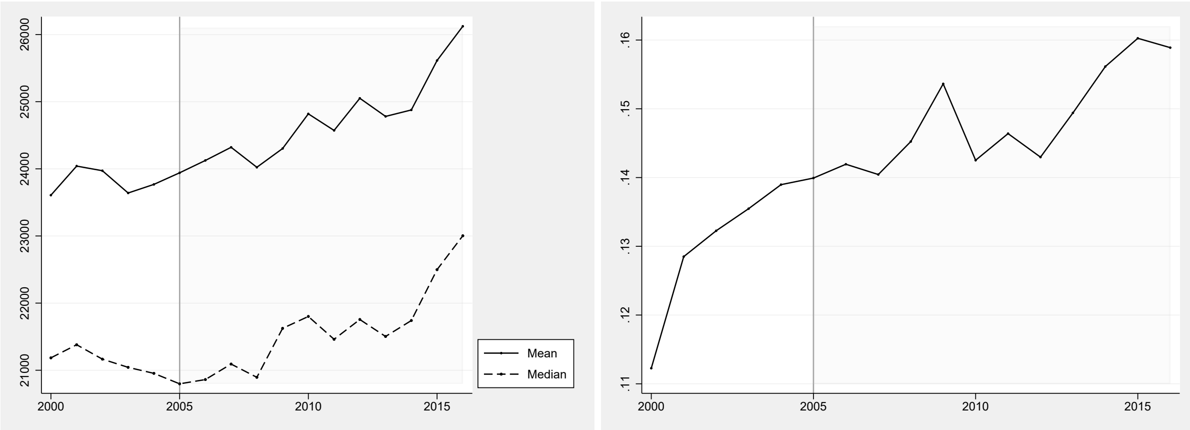


Figure 2.2 – Inequality in equivalised net incomes
(Source: Socio-Economic Panel, own calculations)

The development of mean and median equivalised income is shown in figure 2.3a. After years of stagnation between 2000 and 2005, the average living standard started to grow again in the same year as the employment boom began. Figure 2.3b shows the development of the semi-official 'at-risk-of-poverty rate' (the proportion of individuals with incomes below the relative poverty line of 60 percent of the median), suggesting further strong increases in relative poverty risk after 2010.



(a) Mean/median equivalised income (Euros/year) **(b)** Relative poverty rate

Figure 2.3 – Development of mean/median income and relative poverty rate
(Source: Socio-Economic Panel, own calculations)

Finally, figure 2.4 presents a more detailed description of distributional change for our period under investigation. The figure displays the relative change of the percentiles of the distribution of net equivalised incomes between 2005/06 and 2015/16, indicating in which parts of the distribution (real) income growth was largest. It turns out that all parts of the distribution were shifted upwards, but that growth was relatively modest in the lower part (2.5 to 7.5 percent), larger at the very top (around 7.5 percent), and largest in the upper middle part (7.5 to 10 percent), leading to a small increase in inequality between 2005/06 and 2015/16.[7]

[7] See Dustmann et al. (2022) for an analysis of inequality trends based on the alternative dataset *Einkommens- und Verbrauchsstichprobe (EVS)*. Dustmann et al. (2022) also find slightly rising inequality between 2003 and 2013, but the fact that the EVS is only available every four years as well as a number of differences in survey design make it difficult to compare their analysis with our comparison of the years 2005/06 vs. 2015/16.

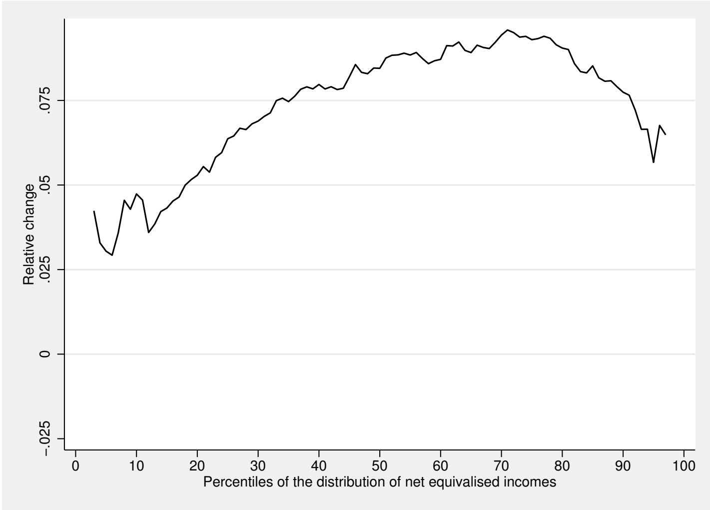


Figure 2.4 – Relative changes of income percentiles of net equivalised income 2005/06 vs. 2015/16
(Source: Socio-Economic Panel, own calculations)

2.4 Empirical analysis

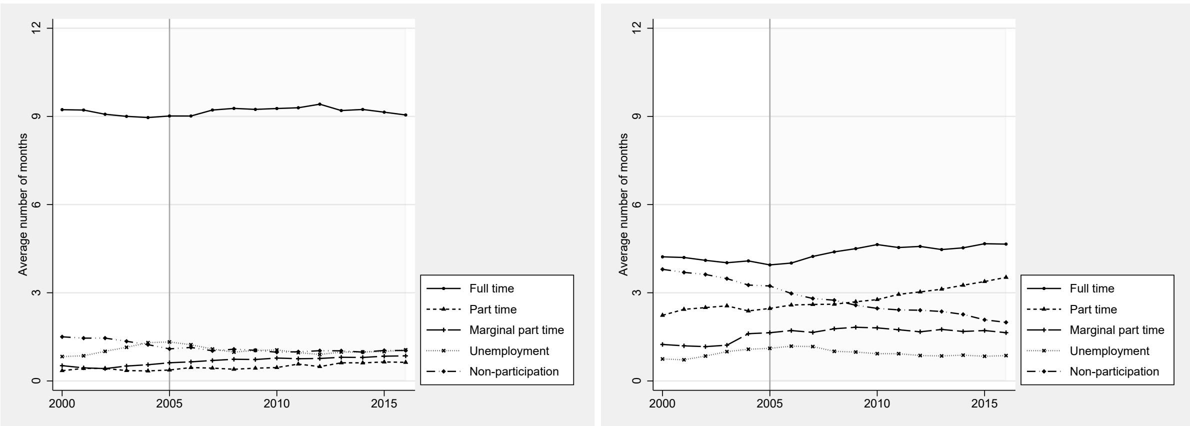
The goal of the following analysis is to determine the contribution of the substantial changes in the level and composition of employment between 2005/06 and 2015/16 with regards to the observed changes in the income distribution, as shown in figure 2.4. In order to assess the role of potential confounders, we also describe the contribution of factors other than employment to the pattern shown in figure 2.4. Our general method will be to compute *counterfactual distributions* of net equivalised incomes for the target period 2015/16 in which we change only one factor (e.g., employment), while keeping all other factors as they are in 2015/16. The comparison of counterfactual vs. factual change will then yield an estimate of the isolated effect of the given factor on the income distribution as observed in 2015/16.[8]

[8] This is commonly accepted methodology in econometric decomposition analysis, see Fortin et al. (2011). It is important to note that this approach does not address general equilibrium effects. On the positive side, it avoids the large number of potentially controversial assumptions that are necessary to model such effects. Policymakers often prefer this approach over equilibrium models for transparency. We view both methods as complementary. The advantage of the method used here is that the potential quantitative importance of different channels of distributional change can be determined in a transparent way with minimal assumptions. It provides an ‘anatomy’ of observed changes that allows us to assess which factors were important and which factors played a negligible role, not claiming their role as final causal determinants. See appendix A.4 for additional discussion.

2.4.1 Distributional effects of the employment boom

Employment changes

We now turn to our analysis of the effects of the employment boom on the distribution of incomes. As described above, our data include detailed information on the annual number of months worked in the different categories of full-time work, part-time work and marginal part-time work. We model counterfactual changes in these employment quantities below. In order to see how the employment boom affected the different forms of employment, we plot in figure 2.5 the evolution of the average number of months worked per year in the different employment categories, separately for men and women. For men, we observe an increase in full-time work between 2005 and 2010, but a stagnation or even decline after 2012 (figure 2.5a). Male part-time and marginal part-time work consistently grew after 2005, albeit at a relatively low level. Interestingly, male non-participation did not decline after 2005 as strongly as unemployment did.



(a) Men **(b)** Women

Figure 2.5 – Annual number of months worked in different employment categories, individuals aged 18–64, not in education
(Source: Socio-Economic Panel, own calculations)

Employment changes were much more dynamic for women (figure 2.5b). In particular, female full-time employment grew considerably between 2005 and 2010. Female part-time employment also consistently increased after 2005, and its growth substantially accelerated after 2010. Female participation in marginal part-time employment also continued to grow after 2005, but growth rates were much lower than in 2003 when this form of employment was liberalised. The steep decline in female non-participation shows that female employment gains mostly came out of non-participation.

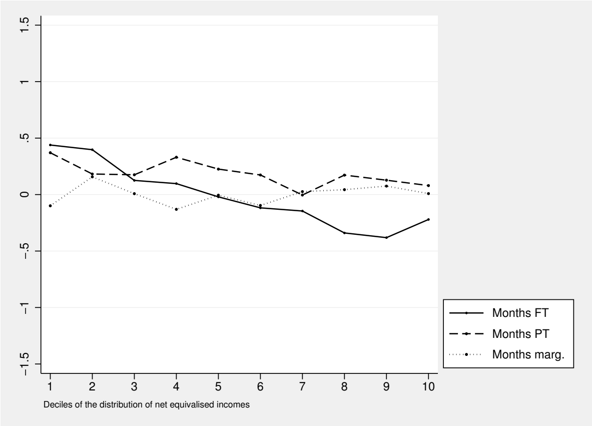
To sum up, the employment boom after 2005 led to substantial employment growth for both men and women, but its most important component was additional full- and part-time employment of women out of non-participation. This evidence is consistent with that from other data sources, see section 2.1, Carrillo-Tudela et al. (2021), Riphahn and Schrader (2020) and appendix A.2.[9]

Before we turn to our detailed micro-analysis of the effect of employment changes on the income distribution, we present a suggestive preliminary analysis aimed at describing the incidence of employment growth across the deciles of the distribution. Figure 2.6a plots the average yearly gains in the number of months worked in the different employment categories per household across different positions of the distribution of equivalised incomes.[10] It appears as if households in the lower part of the distribution substantially gained full-time employment months, while households in the upper part lost over the period 2005–16. This would be a misleading interpretation, however, because it is likely that households in the lower part always tend to gain employment (even in the absence of an employment boom because low income is associated with low employment that can only be increased), while households in the upper part tend to lose employment (because high income is typically associated with a high degree of employment that often cannot be increased further, i.e., mean reversion).

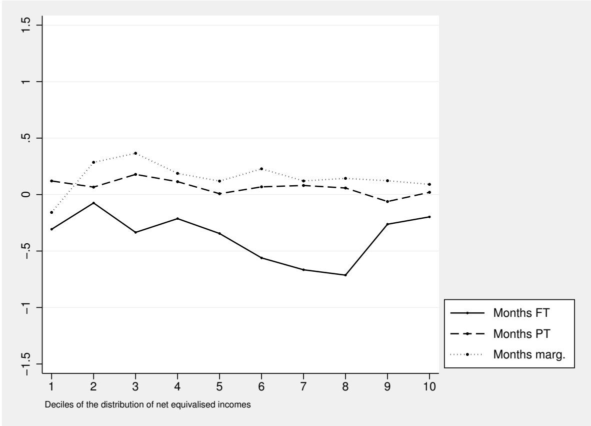
Therefore, in order to determine the effect of the employment boom compared with the previous situation, it makes sense to subtract from figure 2.6a the corresponding figure 2.6b for the period before the boom (i.e., 2000–04). The differential effect shown in figure 2.6c suggests that households at the bottom of the distribution indeed benefitted substantially from full-time employment gains due to the boom and that there were also gains in the middle of the distribution, albeit to a lesser extent. The pattern for part-time employment is similar but not as pronounced. Note that the general level of part-time employment is lower, so that relative gains are still substantial. The growth of marginal part-time employment tends to be negative relative to the period 2000–04. This can be explained by the fact that this type of employment experienced idiosyncratic gains in

[9] Carrillo-Tudela et al. (2021) focus on the age group 25–54 years. Riphahn and Schrader (2020) show that the group of 55–64 year old workers significantly increased their participation after 2005, countering the trend of declining male participation in younger age groups.

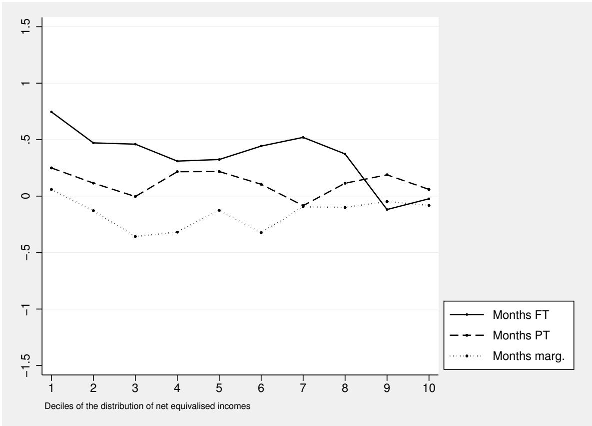
[10] More precisely, we compute for each household from a particular income decile the change in months worked in the different categories from year t to year $t+1$ and average these changes over years and over households from the respective decile.



(a) Average annual growth of the number of months worked per household per year, 2005–16



(b) Average annual growth of the number of months worked per household per year, 2000–04

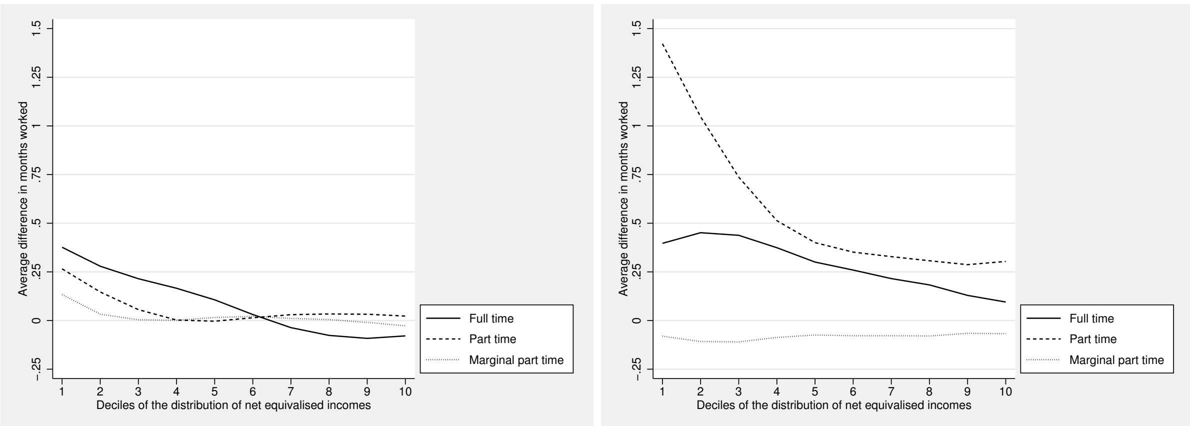


(c) Differential effect (growth incidence in months worked per year 2005–16 relative to 2000–04)

Figure 2.6 – Growth of the absolute annual number of months worked per household across the deciles of the distribution of net equivalised incomes
(Source: Socio-Economic Panel, own calculations)

the year of its liberalisation 2003. Summing up, our preliminary analysis suggests that the employment boom led to employment gains for most parts of the distribution, but that the lower part gained more than the upper part.

We now turn to our more detailed analysis of the effects of the employment boom on the distribution of incomes. Our goal is to model for each individual aged 18–64 and not in education counterfactual employment quantities for 2015/16 that would have prevailed if the boom had not taken place, i.e., if the labour market situation in 2015/16 had been as unfavourable as in 2005/06. In order to do this, we describe the number of months worked per year in the different employment categories (full-time, part-time, marginal part-time) conditional on individual characteristics using logit models.[11] We estimate separate models for each gender and each employment category conditional on the following covariates: nationality, East German residence, disability status, age, age squared, educational qualifications in three categories, work experience, work experience squared and the number of children in different age categories. To account for state dependence in labour market participation, we also include the number of months unemployed/employed in the different employment categories (full-time, part-time, marginal part-time) in the past three years. We estimate such models both for the labour market situation in 2005/06 and 2015/16. Comparing the predictions from these models for a given individual yields a correction term, reflecting how much less/more this individual would have worked in 2015/16 if the labour market situation had still been as in 2005/06. We use this correction term to adjust the factual number of months of each individual observed in 2015/16 into the direction of a counterfactual representing the number of months this individual would have worked in 2015/16 if the employment boom had not taken place.



(a) Men **(b)** Women

Figure 2.7 – Difference in annual months worked 2015/16 compared to counterfactual situation with individual employment probabilities as in 2005/06, individuals aged 18–64, not in education
(Source: Socio-Economic Panel, own calculations)

Figure 2.7 shows the value of these correction terms across the deciles of the distribution of equivalised incomes, separately for men and women.[12] Again, it turns out that absolute employment gains compared to a counterfactual situation in which employment

[11] What follows is an abbreviated description of our calculations. See appendix A for more information and detailed estimation results.

[12] We thank an anonymous referee for suggesting this representation.

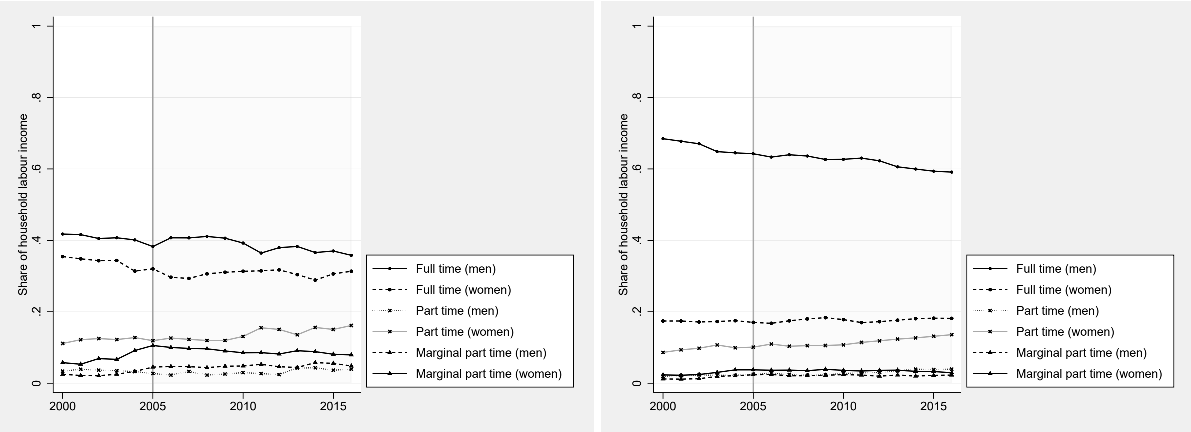
quantities were as in the pre-boom situation of 2005/06 were particularly strong for individuals in the lower part of the distribution. For example, men's average number of full-time months was around a quarter of a month higher in the lower third of the distribution compared to the counterfactual pre-boom situation (figure 2.7a). For women, the average number of part-time months was 0.5 to 1.5 months higher in the lower part of the distribution compared to the pre-boom situation, while the number of full-time months was higher by 0.25 to .5 months. For women, we also observe sizeable gains in part-time and full-time months in the upper half of the distribution (figure 2.7b).

Given the finding that women, especially, increased their participation, it is interesting to see whether this was an added worker effect (i.e., wives entering full-time or part-time work) or increased participation in single households. This can be answered using transition rates between different labour market states (tables A.1 and A.2 in appendix section A.1). The comparison of average annual transition rates for the pre-boom period 2000–04 with those for the period 2005–16 suggests that, for men, the moderate increases in full-time and part-time employment were mainly fuelled by increased transitions from non-participation or part-time employment to full-time employment as well as by increased transitions from unemployment to part-time employment in non-single households. A further striking finding is the decline in the transition rate from unemployment, part-time and marginal part-time into non-participation.[13]

For women, the comparison suggests that the increases in female full-time and part-time work were related to increased transitions from non-participation to full-time or part-time work in non-single households (supporting an added worker effect). Transitions to marginal part-time out of unemployment and non-participation also increased for women in non-single households, while downward transitions from part-time work to marginal part-time also became less frequent. Interestingly, we find *no evidence* for upgrading from part-time to full-time work, neither for single nor for non-single women. While increased transition intensities to full-time and part-time work mostly applied to women in non-single households, the last columns of tables A.1 and A.2 suggest that women in single households experienced a stabilisation of all forms of labour market participation in the sense that transitions from employment or unemployment to inactivity were considerably reduced (this was much less the case for women in non-single households). This indicates

[13] See Carrillo-Tudela et al. (2021) for related evidence, but for monthly rather than yearly transitions.

that also women in single households benefitted from employment gains.



(a) Single adults (one adult, possibly children)

(b) Non-single households (multiple adults, possibly children)

Figure 2.8 – Share of different types of labour income in overall household income, non-pensioner households
(Source: Socio-Economic Panel, own calculations)

The fact that women experienced much higher employment gains than men is also reflected in the rising contribution of female labour income to overall labour income as shown in figures 2.8a and 2.8b. The figures display the share of household labour income that was contributed by men and women in the different employment categories among single and non-single households (excluding pensioner households, see section 2.2 for the definition of different household types). Among single households, full-time labour income of men and women contributed almost equally to overall household labour income, but while the male full-time income share declined, that of women was stable and the female part-time share strongly increased (figure 2.8a). The contribution of female marginal part-time income also declined in favour of more part-time labour income. We conclude that the substantial gains in female participation in part-time work was not only due to an added worker effect but also applied to a large extent to women in single households (possibly with children).

Among non-single households, increasing part-time participation by women also drove up the female part-time labour income share, providing evidence for an added worker effect. By contrast, the share of female full-time income in overall household income increased only slightly. In figure A.2, we show that there were increases in the female full-time labour share in the upper two thirds of the income distribution and in households with children, but these were counteracted by decreases in the lower third of the distribution. Increases in female part-time labour income shares mostly applied to the lower two-thirds of the distribution and were much weaker in the upper third.[14]

Effects on the distribution of net incomes

In order to trace the consequences of these employment changes for equivalised household incomes, we multiply the counterfactual employment months with the monthly wage of the individual in the respective employment category (if observed), or with a monthly wage that we predict using the same set of individual characteristics as in the models for employment in cases in which we do not observe the individual's wage in the respective category (these cases were rare as most individuals *reduced* their employment in the counterfactual pre-boom scenario).

In cases in which individuals counterfactually *lose* employment (because they would have been unemployed or inactive in the labour market absent the employment boom), we check whether these individuals would be entitled to unemployment benefit I (*ALG I*), which depends on the individual labour market history. In order to account for the fact that labour market histories would have been much less favourable in 2015/16 if the employment boom had not taken place, we counterfactually correct each individual's labour market history to reflect how it would have looked under the labour market conditions of 2005/06 (see appendix A.3 for more details). We then calculate the amount of ALG I based on the corrected labour market histories, and impute this income source to all individuals eligible.

In a next step, we sum up all counterfactual income changes per household and recalculate income tax and social security contributions. If the resulting household net income lies below the household minimum income threshold (*'Hartz IV Regelsatz'*) plus housing costs, the household is entitled to the so-called unemployment benefit II (*ALG II*). In these cases, we compute the exact amount of ALG II (plus housing costs) and replace the net income of the household with this amount. Finally, we equivalise the resulting household net incomes using our equivalence scale.

[14]It is possible that the labour income shares also changed because wage rates for the different forms of employment changed. We show in section 2.4.2, however, that this happened only to a very small extent.

The comparison of the counterfactual income distribution for 2015/16 obtained in this way with the factual distribution of 2015/16 reveals which parts of the distribution gained from the boom in terms of net income and to what extent. The results shown in figure 2.9 (dashed line) indicate that the lower part of the distribution benefitted more from the boom than the upper part, consistent with our preliminary analysis in figures 2.6 and 2.7. An important reason why the effects of the boom are not larger is that the consequences of changing back employment quantities to the level of 2005/06 are considerably alleviated by the social security system. If the labour market situation in 2015/16 had been as bad as in 2005/06, not all the individuals affected would have been without income. Many of them would have been entitled to ALG I or II. In order to assess this aspect, the dotted line in figure 2.9 shows the *gross effect* of the boom, i.e., without assigning unemployment benefits to individuals who counterfactually lost employment in our calculations. As expected, this effect is very substantial.

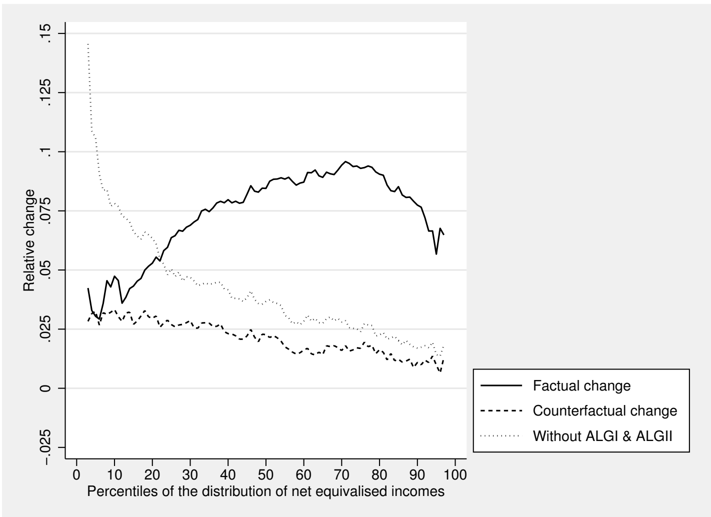


Figure 2.9 – Relative change of income percentiles due to the employment boom
(Source: Socio-Economic Panel, own calculations)

Our counterfactual calculations are supported by figure 2.10 displaying the *factual* changes in the distribution of equivalised incomes *before* and *after* taxes and transfers. Similar to figure 2.9, we observe large relative gains in incomes before taxes and transfers at the bottom of the distribution, which are not translated into corresponding income gains after taxes and transfers. However, figure 2.10 includes the effect of *all other factors* (apart from employment) and does not disentangle the effects of individual aspects as we do in our counterfactual analyses (see below).

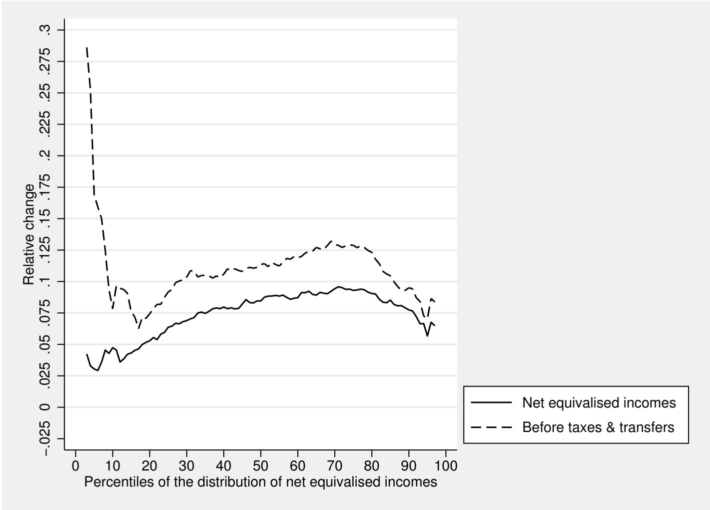


Figure 2.10 – Relative changes of income percentiles 2005/06 to 2015/16 before and after taxes and transfers
(Source: Socio-Economic Panel, own calculations)

Summing up, we draw the following tentative conclusions about the impact of the employment boom on the German income distribution. First, the employment boom led to substantial income gains across the whole distribution. Second, the lower part of the distribution benefitted more than the upper part, most likely because the boom prevented many individuals from being unemployed in 2015/16 (this is implicitly revealed by the difference between the dotted and dashed lines in figure 2.9). Third, the main source of income gains was additional female labour income from part-time and, to a lesser extent, from full-time work out of non-participation or unemployment. Female employment expanded both in single and in non-single households. Fourth, the effects of the boom were substantially dampened by the generous social security system as many of the individuals who gained employment through the boom would have been eligible for unemployment benefits or household minimum income without it. It is well-known that, due to its generosity, in-work net income in the German system is often not much higher than out-of-work net income, especially for households at the bottom of the distribution and/or with many children. In a more general sense, the effects of additional employment on incomes were also dampened by the progressive tax system, which in part taxes away additional income (this applies in particular to additional income earned by second earners in the household due to the joint taxation of spouses in the German tax system). To put it differently, the boom made incomes before taxes and transfers more equal, but this effect was less pronounced after taxes and transfers because the system is effective at counteracting inequality. Fifth, on balance, the boom had an equalising effect on the distribution of net incomes, albeit a moderate one. This follows from column 1 of table 2.3 where we compute the effect of the counterfactual changes on different inequality measures. Finally, sixth, while the boom produced a substantial contribution to overall distributional change (see the solid line in figure 2.9), there must be other factors that also contributed.

2.4.2 Other factors

Given that the employment boom cannot fully account for the changes in the distribution between 2005/06 and 2015/16, we look at a number of other potential explanations: i) immigration, ii) changes in household types, iii) changes in individual and household characteristics, iv) changes in the level and structure of pay, v) changes in capital incomes, and vi) changes in the tax and transfer system. Considering the effect of other factors is important for our understanding of the effects of the boom because its impact may have been wiped out or masked by the countervailing impact of other developments.

Immigration

Like many other countries, Germany experienced substantial immigration during the period under investigation, in particular in the context of the so-called 'refugee crisis' of 2014/15, in the course of which a large number of individuals from the Middle East found refuge in the country. Our database contains information on immigration through a number of refreshment samples (SOEP samples M1 Migration 1995–2010, M2 Migration 2009–13, M3/4 Refugees 2013–15). In order to assess the potential effect of immigration on the distribution of net incomes, we carry out the following counterfactual exercise: we omit all individuals (as well as their children) who immigrated to the country after 2005 from our sample. As in our other computations, this will ignore potential general equilibrium effects of immigration. Such effects are expected to be small, however, as many of the individuals who immigrated after 2005 were refugees who were not allowed to participate in the labour market in the first years after their arrival. Unfortunately, income information on individuals who immigrated as refugees is available for the first time for the year 2016, so that the following results compare 2005/06 to 2016 (rather than to 2015/16).

Migration group	Number of individuals
Aussiedler, Germans living abroad	132,574
EU foreigners	924,646
Asylum seekers/refugees	792,356
Other/no information	1,748,241
Sum	3,598,817

Table 2.1 – Number of individuals who immigrated to Germany between 2005 and 2016
(Source: Socio-Economic Panel, grossed-up numbers using sample weights)

Table 2.1 gives an overview of the number of individuals in our sample counted as having immigrated into the country after 2005 (grossed up to population figures using the sample weights). The total figure of around 3.6 million corresponds well to that reported by the Federal Government (Bundesamt für Migration und Flüchtlinge, 2017). Apart from *Aussiedler* (ethnic Germans born in Eastern European countries with the right to migrate to Germany) and Germans returning from abroad, EU foreigners and refugees constitute the largest groups among the individuals who immigrated after 2005. Our data also contain a large number of immigrants without information on their exact status (the "Other/no information" group in table 2.1). Judged from their observable characteristics, most of these individuals are likely to also belong to the "Asylum seekers/refugees" group.

Figure 2.11 shows the effect of omitting individuals who immigrated since 2005 from the distribution of incomes in 2016. The lower grey line demonstrates that the overall effect of immigration was such that lower parts of the distribution were pulled downwards by up to 4 percent. The other lines show that this was mainly due to the group of refugees and the "no information" group, while the group of EU foreigners and ethnic Germans did not differ much in their composition of incomes compared with the native population. The effect at the lower end is substantial and suggests that the mere fact that a large number of individuals with very low incomes joined the population may account for some of the poor income growth at the bottom of the distribution of net incomes (and may neutralise some of the positive effects of the employment boom). In table 2.3, we show that this had an inequality-increasing effect.

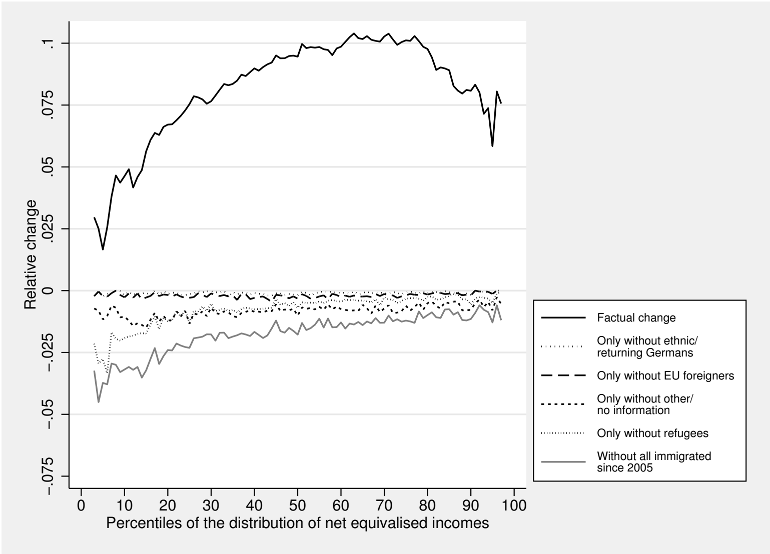


Figure 2.11 – Relative change of income percentiles due to immigration between 2005 and 2016
(Source: Socio-Economic Panel, own calculations)

Changes in household types

Changes in the composition of the population with respect to household types constitute another factor that potentially masks effects of the employment boom on the income distribution. If the share of household types with low equivalent income secularly increases (e.g., lone parents, pensioners), this will lead to increasing inequality independent of employment gains for low-income households.

Figure 2.12 shows that changes in household types over the period under investigation were substantial. In particular, multiple adult households without children and pensioner households increased their population shares at the expense of multiple adult households with children. In order to assess the effect of this development on the income distribution, we counterfactually change the population weights of the different household types in the income distribution of 2015/16 to those in 2005/06 (see appendix A.3 for more details). Figure 2.13a shows that, despite the substantial changes, the effect of doing this is negligible, i.e., changes in household types do not help to account for changes in the distribution between 2005/06 and 2015/16.

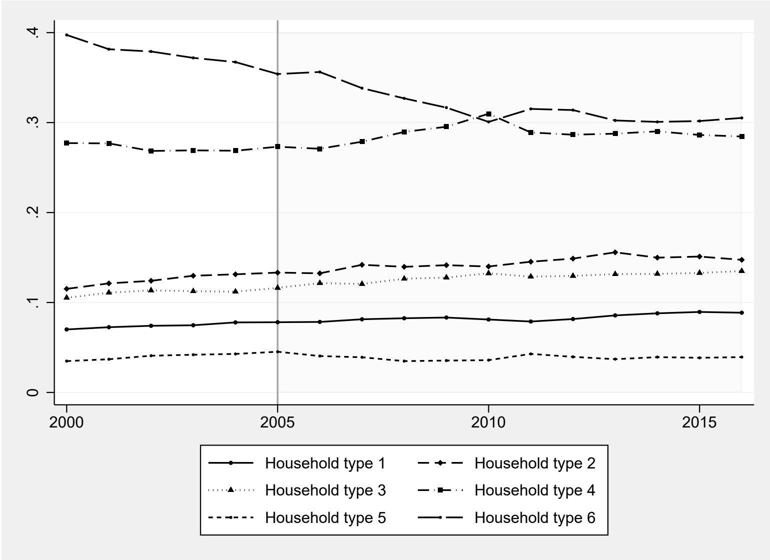
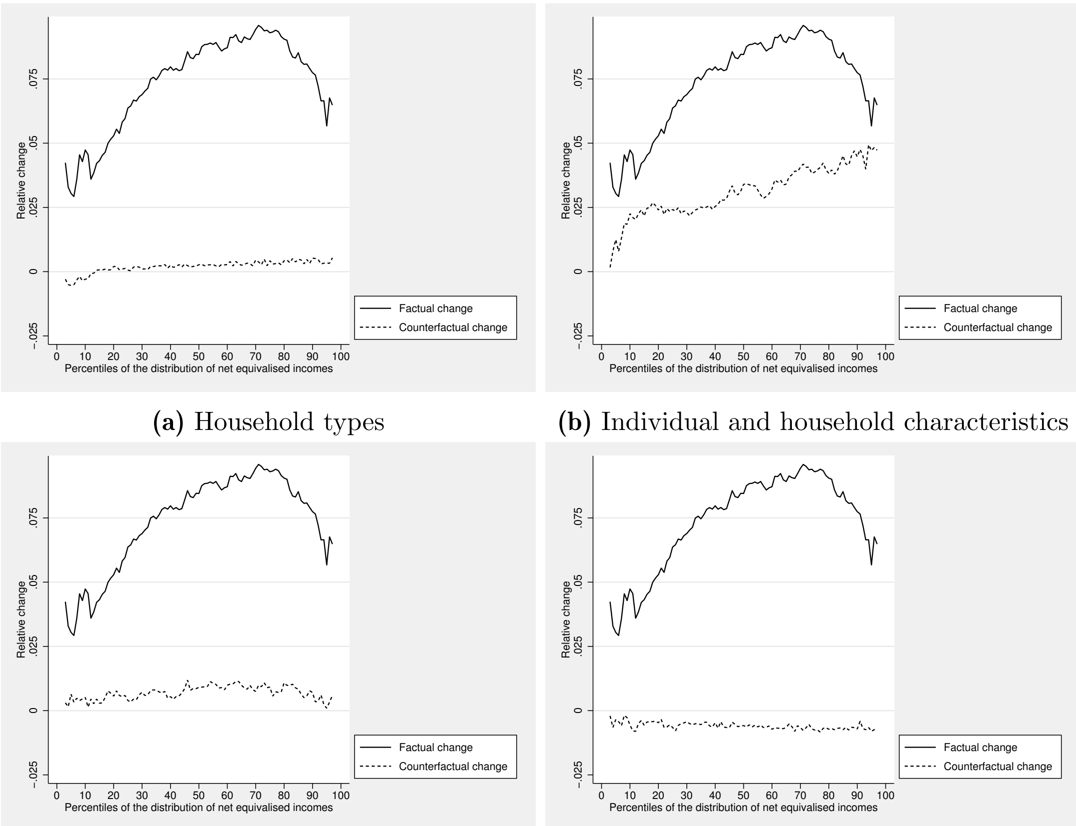


Figure 2.12 – Development of household types over time
(Source: Socio-Economic Panel, own calculations)
Note: household types are 1) single pensioners, 2) multiple pensioners, 3) single adults without children, 4) multiple adults without children, 5) single adults with children, 6) multiple adults with children



(a) Household types

(b) Individual and household characteristics

(c) Level and structure of pay

(d) Capital income

Figure 2.13 – Relative changes of income percentiles due to other factors
(Source: Socio-Economic Panel, own calculations)

Changes in individual and household characteristics

Next, we consider finer compositional changes in the structure of the population. For example, it may be the case that educational upgrading and population ageing induced more income inequality because a shift towards higher educational qualifications and older age groups raised the share of population subgroups with high income dispersion (increasing within-group inequality), or increased the divide between education or age groups (increasing between-group inequality). Table 2.2 summarises the changes in the individual and household characteristics we have considered. As expected, there is a trend towards higher age, work experience and education as well as towards more households with female heads and fewer children.

Variable	Average 2005/06	Average 2015/16	Difference
Household head			
Female	0.37	0.43	+0.06
Foreign nationality	0.08	0.10	+0.02
Age	50.18	52.47	+2.29
University degree	0.19	0.25	+0.06
Vocational training	0.63	0.59	-0.04
Less than vocational training	0.18	0.15	-0.02
Work experience (years)	12.22	13.35	+1.13
Partner or second oldest person (if any)			
Female	0.72	0.63	-0.09
Foreign nationality	0.11	0.10	-0.01
Age	46.97	49.28	+2.31
University degree	0.14	0.17	+0.03
Vocational training	0.59	0.51	-0.08
Less than vocational training	0.27	0.32	+0.04
Work experience (years)	9.95	10.40	+0.45
Other household characteristics			
East Germany	0.21	0.20	-0.01
Number of children in household	0.70	0.61	-0.08
Number of children 0-3 years	0.11	0.11	0.00
Number of children 4-6 years	0.12	0.10	-0.02
Number of children 7-17 years	0.47	0.41	-0.07
More than two adults	0.18	0.17	0.00

Table 2.2 – Individual and household characteristics in 2005/06 and in 2015/16
(Source: Socio-Economic Panel, own calculations)

We compute the effect of these changes on the income distribution in 2015/16 by reweighting the distribution of these characteristics back to the one observed in 2005/06, leaving everything else constant. We do this separately by household type using the semi-parametric reweighting procedure proposed by DiNardo et al. (1996) (see appendix A.3 for more details). Figure 2.13b shows that the impact of changes in these characteristics on the distribution of incomes was considerable. The shift towards higher age and education groups implied higher income levels, especially in the middle and at the top of the distribution. This contributed to increasing inequality, counteracting the pro-poor income growth induced by the employment boom (column 4 of table 2.3). The compositional effects of changing income-relevant characteristics is consistent with findings in the literature showing that these can account for a large part of changes in the distribution of *wages* (Biewen and Seckler, 2019; Dustmann et al., 2009).

Changes in the level and structure of pay

Apart from quantities, prices for employment may have changed over the period we have considered. In order to describe the potential effect of this factor on the distribution of net incomes, for each individual observed in 2015/16 we form a counterfactual wage that mimics the wage this person would have earned under the pay structures of 2005/06.[15] To this end, we regress monthly (log) wages on the following characteristics: nationality, East German residence, disability status, age, age squared, three education categories, work experience and work experience squared. We do this separately for the three employment categories, the two genders, and the two situations 2005/06 and 2015/16, the latter representing the pay structures in 2005/06 and in 2015/16, respectively. We then compute for each individual observed working in 2015/16 a correction term based on the difference in wage predictions under the pay structures of 2005/06 and 2015/16, reflecting how much higher/lower the individual's wage would have been under the pay structure of 2005/06. We also consider changes in pay for unobservables (i.e., wage residuals) assuming that the individual would have had the same rank in the distribution of residual wages in 2005/06 as in 2015/16. The resulting counterfactual wages are then multiplied by the observed number of months worked in the different employment categories, yielding changes in individual and household market income. Finally, we compute taxes and social security contributions for the changed sum of incomes and carry out the equivalisation.

[15] See appendix A for more information and detailed regression results.

Note that this procedure captures changes both in the *level* and in the *structure* of wages.[16] The results of this exercise are shown in figure 2.13c. It turns out that changes in pay played only a minor role for the development of the income distribution between 2005/06 and 2015/16. There were small real wage gains, which were slightly higher for the middle of the distribution. This did not significantly impact income inequality (see lower panel of table 2.3). The (missing) effect of changes in pay structures for the period under investigation found in our analysis is consistent with evidence from administrative data showing that, after increasing inequality before 2005, the quantiles of the wage distribution mostly developed in a horizontal way, implying stagnating real incomes and no increasing inequality after 2005 (see Baumgarten et al., 2020, p. 7, Figure 1b). The important conclusion for our analysis is that the income effects of the boom did not operate much through increasing wages in a tighter labour market or by changes in productivity sharing but mainly by increasing employment quantities at constant pay.

Changes in capital incomes

Changes in capital incomes may also have influenced the income distribution in the period we have considered. We investigate this by constructing a counterfactual distribution of net incomes that results if one changes back the distribution of capital incomes to its state in 2005/06, leaving everything else constant. We do this by transforming each household's rental income and each household's other capital incomes by multiplying them by the ratio of the percentiles of these distributions in 2005/06 and 2015/16 based on the corresponding ranks of the household in 2015/16 (see appendix A.3 for more details). Again, this reflects both changes in the *level* and the *dispersion* of capital incomes.

The effect of changing rental and other capital incomes is shown in figure 2.13d. The figure suggests that changes in capital incomes *depressed* the income distribution. This is in line with the fact that real interest rates fell over the period considered. Perhaps surprisingly, these effects occured uniformly across the distribution. Our analysis comes with the caveat that survey data (e.g., from SOEP) do not cover developments at the very top of the income distribution (Bartels and Jenderny, 2015). However, the results in

[16] Recall that we only consider changes in real wages as all of our wage information is expressed in prices of 2016.

Bartels and Jenderny (2015) suggest that changes at the very top of the German income distribution were relatively modest compared to those in other countries such as the United States. Also note that the respondents in our survey may report certain capital incomes as income from self-employment (in our study included in labour income). Drechsel-Grau et al. (2015) have shown on the basis of tax data that, if one excludes incomes from owner-run enterprises, capital incomes are indeed approximately uniform across the German income distribution. Overall, we do not find any evidence for an important role of capital incomes for changes in the distribution of net incomes, but certainly cannot rule out effects at the very top not covered by our data.

Changes in the tax and transfer system

We consider the effect of the following changes in the German tax and transfer system that occurred between 2005/06 and 2015/16.

Changes in transfers:

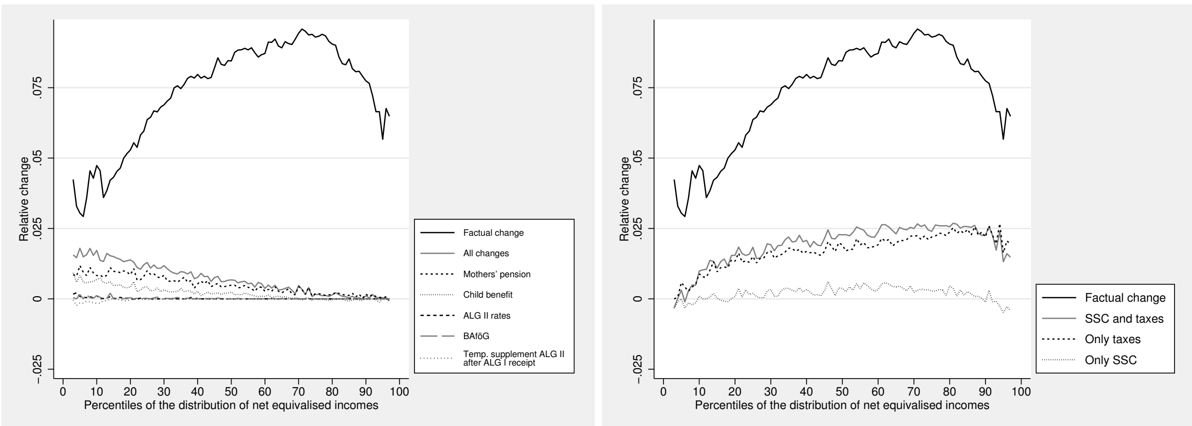
- Extension of mothers' pensions (two, instead of one year, of implicit contributions for children born before 1992)
- Abolishment of the temporary supplement to ALG II after receipt of ALG I (transitionary payment for individuals whose ALG I ran out amounting to 2/3 of the difference between ALG I and ALG II in the first year, and 1/3 in the second year)
- Higher child allowances, higher student allowances, higher ALG II (we only consider the part of the increase since 2005 that was higher than inflation)

Changes in the tax system (including changes in social security contribution rates):

- Introduction of a 'rich tax' (marginal tax rate of 45 percent instead of 42 percent starting from 250,000 (500,000) euros taxable income per annum)
- Withholding tax for capital incomes (flat rate of 25 percent instead of personal tax rate)
- Changes in the tax schedule (changes in a number of tax allowances plus various changes in marginal tax rates)

- Changes in social security contribution rates (mainly reductions, e.g., lower contribution rates to unemployment insurance due to falling unemployment)

We describe the effects of these changes on the distribution of net incomes by counterfactually undoing each of these reforms. We emphasise that, as in our other computations, we ignore potential behavioural reactions to these changes.[17] The results of these operations are shown in figures 2.14a and 2.14b. Figure 2.14a and the numbers in table 2.3 demonstrate that the changes in the transfer system tended to have an equalising effect, mainly due to the extended mothers' pension and the higher child allowances. However, the changes in the tax schedule mainly benefitted households in the middle and the top part of the distribution (households at the bottom of the distribution typically do not pay income tax). This had an inequality-increasing effect (figure 2.14b and table 2.3). The fall in social security contribution rates led to small income gains in the middle of the distribution, but not at the bottom and the top (households at the bottom are typically not employed and labour incomes in households at the top typically exceed the social security contributions ceiling).



(a) Transfer system

(b) Tax and SSC system

Figure 2.14 – Relative change of income percentiles due to tax and transfer changes
(Source: Socio-Economic Panel, own calculations)

[17]Such reactions are likely to be small and they typically counteract the original effects (Jessen, 2019), rendering our calculations upper bounds.

Index	Employment boom	Immigration	HH types	HH characteristics
Mean	+471.037	-362.008	+79.502	+908.138
	[+314.302 ; +626.474]	[-429.526 ; -293.971]	[-49.634 ; +205.381]	[+668.401 ; +1146.968]
Median	+512.069	-366.039	+61.764	+749.205
	[+320.035 ; +717.644]	[-485.849 ; -254.829]	[-62.887 ; +188.914]	[+510.938 ; +972.845]
P90/P10	-0.073	+0.078	+0.025	+0.090
	[-0.131 ; -0.019]	[+0.038 ; +0.121]	[+0.002 ; +0.053]	[+0.021 ; +0.156]
P90/P50	-.019	+0.011	+0.004	+0.025
	[-0.036 ; +0.000]	[-0.004 ; +0.024]	[-0.004 ; +0.014]	[-0.001 ; +0.051]
P50/P10	-0.019	+0.031	+0.010	+0.023
	[-0.047 ; +0.006]	[+0.010 ; +0.052]	[+0.000 ; +0.021]	[-0.003 ; +0.048]
Gini	-0.003	+0.004	+0.001	+0.005
	[-0.005 ; -0.001]	[+0.003 ; +0.005]	[+0.000 ; +0.002]	[+0.002 ; +0.008]
Poverty rate	-0.002	+0.005	+0.001	+0.004
	[-0.007 ; +0.002]	[+0.002 ; +0.008]	[-0.001 ; +0.003]	[-0.001 ; +0.009]

Index	Pay structures	Capital incomes	Transfers	Tax and SSC
Mean	+160.299	-159.121	+117.167	+526.212
	[-63.717 ; +385.391]	[-210.710 ; -108.257]	[+112.257 ; +121.559]	[+506.540 ; +544.972]
Median	+209.604	-144.170	+154.761	+518.147
	[+26.584 ; +400.229]	[-198.236 ; -87.573]	[+117.870 ; +196.741]	[+452.883 ; +585.615]
P90/P10	+0.001	-0.006	-0.053	+0.050
	[-0.055 ; +0.057]	[-0.026 ; +0.014]	[-0.074 ; -0.030]	[+0.024 ; +0.080]
P90/P50	-0.005	-0.002	-0.012	+0.001
	[-0.030 ; +0.018]	[-0.009 ; 0.006]	[-0.016 ; -0.007]	[-0.012 ; +0.014]
P50/P10	+0.006	-0.001	-0.016	+0.027
	[-0.012 ; +0.025]	[-0.010 ; +0.006]	[-0.027 ; -0.004]	[+0.016 ; +0.038]
Gini	-0.001	-0.000	-0.003	+0.001
	[-0.004 ; +0.004]	[-0.001 ; +0.000]	[-0.003 ; -0.002]	[+0.001 ; +0.002]
Poverty rate	+0.002	-0.000	-0.003	+0.004
	[-0.001 ; +0.006]	[-0.002 ; +0.001]	[-0.005 ; -0.002]	[+0.002 ; +0.007]

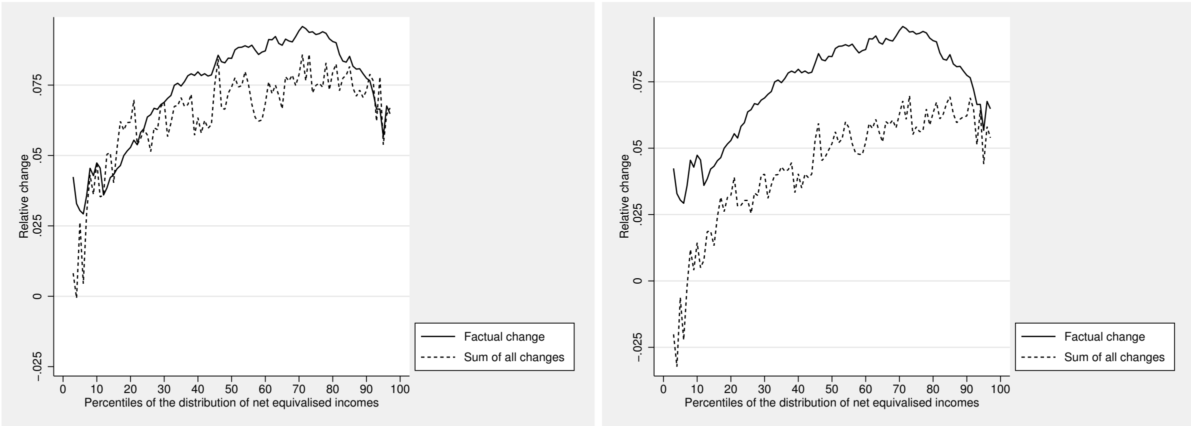
Index	Sum	Factual change
Mean	+1741.358	+1919.746
	[+1322.354 ; +2162.391]	[+1477.581 ; +2344.251]
Median	+1695.39	+1921.334
	[+1244.221 ; +2154.239]	[+1539.278 ; +2312.802]
P90/P10	+0.112	+0.121
	[-0.027 ; +0.261]	[-0.011 ; +0.246]
P90/P50	+0.003	-0.014
	[-0.057 ; +0.061]	[-0.063 ; +0.038]
P50/P10	+ 0.059	+0.080
	[-0.001 ; +0.117]	[+0.027 ; +0.136]
Gini	+0.005	+0.003
	[-0.001 ; +0.011]	[-0.006 ; +0.011]
Poverty rate	+0.011	+0.020
	[-0.001 ; +0.022]	[+0.011 ; 0.030]

95% bootstrap confidence intervals in brackets (1000 replications)

Table 2.3 – Effects on inequality measures

(Source: Socio-Economic Panel, own calculations)

2.4.3 Summary of changes



(a) With employment boom

(b) Without employment boom

Figure 2.15 – Factual change vs. sum of counterfactual changes with and without employment boom
(Source: Socio-Economic Panel, own calculations)

How successful are our calculations at putting together Germany's inequality puzzle? Figure 2.15a shows that the sum of all changes modelled by us reconstructs the observed changes in the distribution strikingly well. This is also the case for the inequality calculations in table 2.3, although these are more affected by the non-smooth form of the sum of changes (figure 2.15a). Figure 2.15b, showing the sum of all changes *without* the employment boom, suggests that the employment boom indeed contributed substantially to the distributional changes between 2005/06 and 2015/16, but that its impact was masked by a number of other developments, which also undid some of its inequality-reducing effects.

2.5 Conclusion

In this paper, we have addressed the question of why income inequality and poverty risk remained persistently high in Germany, despite an unprecedented labour market boom that drastically reduced unemployment and particularly boosted employment in the lower part of the distribution. We reach the following conclusions. First, the boom indeed increased incomes and reduced inequality as lower parts of the distribution benefitted more than the middle or the upper part. Second, the effects of the boom on net incomes were substantially dampened by the social security system and the progressive tax system, which reduce the impact of economic downturns on disposable incomes, but also that of economic upturns. Third, much of the boom took the form of additional female part-time and full-time employment, demonstrating that such employment gains may not only boost incomes but have equalising effects. Fourth, the effects of the boom on the income distribution were masked by a number of other developments such as immigration of individuals with low incomes and changes in the composition of the population (educational upgrading and population ageing), making it difficult to determine their exact magnitude. Finally, our results imply that if the COVID-19 shock were to reverse all the employment gains that occurred during the boom, this would only have a moderately disequalising effect on the distribution of net incomes due to the strong role of the German tax and transfer system as a distributional stabiliser.

Appendix A

A.1 Transition rates between employment states

In the following tables, the main yearly labour market state is defined as the one with the majority of months. Ties were resolved according the following ordering: full-time, part-time, marginal part-time, unemployment, non-participation.

		Men					Women					
		FT	PT	MPT	UNEM	NP		FT	PT	MPT	UNEM	NP
Single	FT	0.920	0.019	0.005	0.038	0.015	FT	0.876	0.039	0.011	0.029	0.043
	PT	0.231	0.620	0.043	0.046	0.060	PT	0.145	0.672	0.064	0.048	0.065
	MPT	0.188	0.104	0.404	0.134	0.162	MPT	0.069	0.125	0.570	0.062	0.155
	UNEM	0.142	0.033	0.046	0.650	0.128	UNEM	0.109	0.069	0.073	0.617	0.131
	NP	0.083	0.016	0.068	0.074	0.662	NP	0.058	0.039	0.044	0.060	0.701
		FT	PT	MPT	UNEM	NP		FT	PT	MPT	UNEM	NP
Non-single	FT	0.926	0.012	0.008	0.027	0.023	FT	0.830	0.069	0.018	0.030	0.053
	PT	0.248	0.403	0.133	0.059	0.151	PT	0.100	0.741	0.078	0.024	0.055
	MPT	0.152	0.080	0.448	0.080	0.220	MPT	0.051	0.149	0.619	0.027	0.151
	UNEM	0.199	0.018	0.048	0.579	0.154	UNEM	0.112	0.065	0.063	0.562	0.198
	NP	0.118	0.031	0.077	0.038	0.654	NP	0.047	0.059	0.066	0.024	0.783

Table A.1 – Average yearly transition rates 2000-2004

Average yearly transition rates between employment states (origin = rows, destination = columns) based on main employment state in given year (majority of months), 2000-2004, individuals aged 18 to 64 years, not in education (Source: Socio-Economic Panel, own calculations)

		Men					Women					
		FT	PT	MPT	UNEM	NP		FT	PT	MPT	UNEM	NP
Single	FT	0.936	0.014	0.010	0.021	0.013	FT	0.900	0.040	0.014	0.014	0.026
	PT	0.234	0.571	0.043	0.079	0.068	PT	0.149	0.717	0.043	0.033	0.044
	MPT	0.120	0.096	0.570	0.128	0.080	MPT	0.098	0.127	0.549	0.100	0.106
	UNEM	0.127	0.029	0.066	0.693	0.074	UNEM	0.077	0.068	0.100	0.659	0.087
	NP	0.078	0.020	0.084	0.060	0.681	NP	0.066	0.053	0.081	0.061	0.658
		FT	PT	MPT	UNEM	NP		FT	PT	MPT	UNEM	NP
Non-single	FT	0.938	0.014	0.008	0.015	0.018	FT	0.849	0.071	0.015	0.011	0.053
	PT	0.302	0.502	0.077	0.029	0.082	PT	0.095	0.806	0.038	0.016	0.043
	MPT	0.151	0.067	0.488	0.077	0.198	MPT	0.053	0.109	0.671	0.025	0.139
	UNEM	0.199	0.037	0.081	0.561	0.116	UNEM	0.090	0.073	0.097	0.571	0.167
	NP	0.137	0.034	0.111	0.031	0.628	NP	0.062	0.072	0.102	0.026	0.721

Table A.2 – Average yearly transition rates 2015-2016

Average yearly transition rates between employment states (origin = rows, destination = columns) based on main employment state in given year (majority of months), 2015-2016, individuals aged 18 to 64 years, not in education (Source: Socio-Economic Panel, own calculations)

A.2 Employment information in the SOEP

As described in the main text, we use information on the annual number of months worked in different employment categories along with their associated earnings in order to carry out counterfactual analyses. The information we use comes from two parts of the survey: (i) summary information on annual number of months spent in full-time, part-time and marginal part-time work (employment calendar), (ii) summary information about the earnings from different sources: self-employment, main job and side job, along with the annual number of months worked for these categories (income calendar). The information in both calendars is in most cases consistent, but in a number of cases we observe employment spells or earnings that appear in one calendar but not in the other. In order to derive variables for the number of months worked in full-time, part-time and marginal part-time that consistently add up with the reported earnings to the total annual labour income of the individual (resulting from the income calendar), we make certain adjustments for these cases. For example, if we observe information in the income calendar but no associated spells in the employment calendar, we impute these spells into the other calendar using information on the level of pay along with other information and vice versa (codes are available on request).

Despite limitations with regard to comparability, the employment numbers in the SOEP generally correspond quite well to those from official sources, see table A.3. A notable exception are the numbers for dependent full-time employment for which our SOEP data generally implies both higher employment rates and for men also smaller increases due to the boom. One likely explanation for these differences is the fact that the employment numbers provided by the Federal Employment Agency only represent officially reported employment subject to social security contributions. Thus, they exclude substantial parts of overall employment such as civil servants (not subject to social security contributions), cross-border employment and other forms of potentially unreported employment such as work in family businesses or in the shadow economy. As to the latter, it is likely that a number of jobs that went unreported before the boom were transformed into officially reported jobs subject to social security contributions during the more stable boom conditions. Another potential difference is the sometimes unclear distinction between self-employment and dependent employment. The fact that SOEP numbers for dependent employment are generally higher but those for self-employment lower compared to the other sources suggests that SOEP participants may report jobs as dependent full-time employment although they could be defined as self-employment.

	Men			
	SOEP		Official sources	
Employment rate	2005/06	2015/16	2005/06	2015/16
Self-employed	0.091	0.084	0.104	0.102
Dependent full-time work	0.615	0.626	0.519	0.575
Part-time	0.031	0.054	0.028	0.065
Marginal part-time (main job)	0.098	0.106	0.094	0.117
Marginal part-time (side job)	0.028	0.048	0.031	0.050
	Women			
	SOEP		Official sources	
	2005/06	2015/16	2005/06	2015/16
Self-employed	0.029	0.037	0.055	0.058
Dependent full-time work	0.302	0.342	0.283	0.306
Part-time	0.194	0.267	0.181	0.264
Marginal part-time (main job)	0.179	0.178	0.169	0.184
Marginal part-time (side job)	0.028	0.048	0.031	0.050

Table A.3 – Employment rates

Employment rates in population aged 18-64 according to SOEP and official sources (Federal Statistical Office, Federal Employment Agency, Eurostat)

SOEP: An individual was classified as self-employed, full-time employed, part-time employed if this activity represented the majority of months in the employment calendar. An individual was classified having marginal employment if she reported a positive number of months employed in this category. An individual was classified as having marginal employment as a side job if she reported a positive number of months employed in marginal employment but was mainly classified as self-employed, full-time or part-time employed.

Official sources: All employment rates were calculated from absolute numbers which were divided by the size of the population aged 18-64 years (Source: Federal Statistical Office). The information on full-time and part-time employment stems from the Federal Employment Agency (administrative social security data) and refers to the number of jobs subject to social security contributions at the 30th June of each year. We corrected the structural break in reporting full-time vs. part-time employment from 2010 to 2011 (Fitzenberger and Seidlitz, 2020) by adding the jump in part-time employment between 2010 and 2011 to the numbers for part-time employment before 2011 and by subtracting it from the numbers of full-time employment before 2011. The numbers on marginal part-time jobs are also due to the Federal Employment Agency. For marginal part-time jobs in addition to other employment, no separate numbers are available for men and women (we divided these numbers proportionally among men and women). The numbers on self-employment (not included in the data of the Federal Employment Agency) are from the the Labor Force Survey/Eurostat as reported in Bonin et al. (2020).

A.3 Counterfactual analysis

This section documents more details of our counterfactual analyses.

Employment boom

The distribution of months worked per year in the different employment categories turns out to be very bipolar with little mass on the intermediate outcomes 1 to 11 months. Preliminary experiments with different ordinal models suggested that it is practically impossible to predict the exact number of months outside 0 or 12 months, or even whether this number lies between 0 and 12. Therefore, and in order to preserve the clear distinction between participation and non-participation, we defined in our logit models for each employment category an outcome zero (0 months worked per year in the respective employment category) and an outcome one (1 to 12 months worked per year). In order to eventually also obtain predicted values 1 to 11 for months worked per year as observed in the data, we calculated for each observation a "rounding correction" which is equal to the observed value minus the "rounded value" (the rounded value is zero months for actual zero months, and twelve months for actual 1 to 12 months). The implied rounding correction is zero for observed months zero, and equal to the observed value minus twelve for all other observations. We then predict with our logit model zeros and ones in the usual way using a probability cutoff and assign to logit predictions zero the value zero months and to logit predictions one initially the value twelve months. We then add back the "rounding correction", capping all resulting values at zero months (if negative) or twelve months (if larger than twelve). In this way, we obtain a distribution of predicted months that very closely resembles the distribution of observed months and that at the same time preserves for a large number of observations the fact that the corresponding individual had an intermediate value of months worked per year.

In a robustness analysis illustrated in appendix A.6, we vary the definition of outcomes in our logit models as follows: (i) Categorisation of 0 to 5 months into zero and of 6 to 12 months into one, applying an analogous rounding correction to predicted values zero and twelve months (i.e., adding back the difference between the observed values and "rounded values" zero and twelve, always capping predictions at zero if negative and at twelve if larger than twelve). (ii) Categorisation of 0 months into zero and of 1 to

12 months into one, but no rounding correction (i.e., predictions are always either zero or twelve months). For robustness analysis (i) the results of which are shown in figure A.3, we obtain slightly larger income gains in the lower part of the distribution, while robustness analysis (ii) yields almost identical results to our main analysis, see figure A.4.

In order to compute our counterfactual predictions, we estimate logit models[18] for the months worked (separately for full-time, part-time, marginal part-time) in the two labour market situations 2005/06 (period 0) and 2015/16 (period 1), conditional on the following individual characteristics: nationality, East German residence, disability status, age, age squared, three education categories, work experience, work experience squared, the number of children in different age categories as well as the number of months unemployed and employed in the last three years in the three employment categories. We estimate these models separately by gender and employment category. From each logit model we calculate the number of predicted months in each employment category as described in the first paragraph of this section. We use the difference in predicted number of months employed in each category for 2005/06 vs. 2015/16 per individual to correct the actual number of months worked by individuals observed in 2015/16 into the direction of the labour market situation of 2005/06 (i.e., before the boom):[19]

$$MonthsFT_i^{1,cf} = MonthsFT_i^1 + (\widehat{MonthsFT}_i^0 - \widehat{MonthsFT}_i^1) \tag{A.1}$$
$$MonthsPT_i^{1,cf} = MonthsPT_i^1 + (\widehat{MonthsPT}_i^0 - \widehat{MonthsPT}_i^1) \tag{A.2}$$
$$MonthsMarg_i^{1,cf} = MonthsMarg_i^1 + (\widehat{MonthsMarg}_i^0 - \widehat{MonthsMarg}_i^1). \tag{A.3}$$

The correction terms reflect how much lower/higher the number of months worked by the individual would have been if the labour market situation in 2015/16 had still been as in 2005/06, given her observed characteristics. We cap the counterfactually predicted number of months worked $MonthsFT_i^{1,cf}$, $MonthsPT_i^{1,cf}$, $MonthsMarg_i^{1,cf}$ at 0 or 12 in case they lie outside the interval 0 to 12. To account for the connectedness of decisions

[18]Full estimation results are contained in appendix A.7, see tables A.4 to A.6

[19]For the terms in brackets, note that the "rounding correction" applied to compute the predictions is the same for 2005/06 and 2015/16, i.e., it does not include changes in the prevalence of intermediate values 1-11 months worked. However, the proportion of such intermediate values is small (lower than 10% or lower than 5% depending on the employment category) and it does not change much over time. As a consequence, results are not much influenced by this aspect.

in the three employment categories, we include the number of full-time months as a regressor in the models for part-time months and the number of full-time/part-time months in the models for marginal part-time. We also cap the counterfactual number of part-time months $MonthsPT_i^{1,cf}$ at twelve minus the counterfactual number of full-time months in order to rule out that the combined counterfactual number of full-time and part-time months exceeds twelve. We do not cap the counterfactual number of marginal part-time months $MonthsMarg_i^{1,cf}$ (other than constraining it to lie in the interval 0 to 12) in order to allow for the possibility of fully parallel spells of marginal part-time along full- or part-time employment (such cases exist in our analysis as well as in the real world).

For our employment models, we only consider individuals aged between 18 and 64 years who are not in education. Our logit predictions are specified such that the counterfactual distribution of months worked in the different categories resembles the factual distribution of 2005/06. In order to take account of the fact that labour market histories in the situation of 2005/06 would generally have been much less favourable than in the situation of 2015/16, we also correct the observed number of months unemployed/worked in the past three years (separately by full-time, part-time, marginal part-time status) for each individual before we compute counterfactual predictions using a similar procedure as for the number of months worked (i.e., we estimate regressions for these quantities both for 2005/06 and 2015/16 and use the difference in predictions to correct the values observed for 2015/16 into the direction of 2005/06).

The counterfactual annual labour income of individuals observed in 2015/16 reflecting the labour market situation of 2005/06 is then computed as

$$MonthsFT_i^{1,cf} \cdot \overline{WageFT}_i^1 + MonthsPT_i^{1,cf} \cdot \overline{WagePT}_i^1 + MonthsMarg_i^{1,cf} \cdot \overline{WageMarg}_i^1 \tag{A.4}$$

where $\overline{WageFT}_i^1$, $\overline{WagePT}_i^1$ and $\overline{WageMarg}_i^1$ denote the monthly wages of the individual in the respective employment category. If the monthly wage of the individual in an employment category with non-zero counterfactual months is not observed, we predict it based on wage regressions conditional on the following individual characteristics (separately by gender): nationality, East German residence, disability status, age, age squared, three education categories, experience and experience squared.[20]

If an individual is hit by a counterfactual loss of at least six full-time months (relative to the observed number of months worked in 2015/16), we check whether this individual would be entitled to unemployment benefit I (ALG I). For this, we use the employment history of the individual in the past three years which was corrected earlier for the fact that employment histories in 2005/06 were less favourable than 2015/16 (see above). If the individual is entitled to unemployment benefit I in the counterfactual state, we compute the exact entitlement per individual and month and assign it to the individual for the number of counterfactually lost employment months.

In the next step, we sum up all income sources per household (including the counterfactually changed labour incomes) and recompute taxes and social security contributions. The resulting counterfactual household net income is given by (in simplified notation):

$$\begin{aligned} y_{cf} = {} & y_{Market} + \hat{\Delta} y_{Labour} + y_{Pens} + y_{Trans} + \hat{\Delta} y_{Trans} \\ & - ssc_1(y_{Labour} + \hat{\Delta} y_{Labour}, y_{Pens}) - tax_1(y_{Tax} + \hat{\Delta} y_{Labour}), \end{aligned} \tag{A.5}$$

where y_{Market}, y_{Labour}, y_{Pens}, y_{Trans} and y_{Tax} denote household market income, household labour income, household pension income, transfers received by the household and the household's taxable income, respectively. The terms $\hat{\Delta} y_{Labour}$ and $\hat{\Delta} y_{Trans}$ incorporate the counterfactual changes in household labour incomes and the counterfactual addition/subtraction of ALG I due to losses/gains in employment. The changes in labour incomes further feed into social security contributions and taxes, as reflected by the last two components in equation (A.5).

In a last step, we check whether the above net household income falls below the subsistence level of unemployment benefit II (ALG II or *Hartz IV*) plus costs for accommodation and heating. If this is the case, y_{cf} is replaced by the latter.

Immigration

To assess the impact of immigration on the distribution of net incomes between 2005/06 and 2016, we omit for the year 2016 all individuals who immigrated into the country

[20]These are the same wage regressions as in section A.3 and are reported in appendix A.7, see tables A.9 to A.11.

between 2005 and 2016 as well as children below 16 years of age living in their households (see main text for more details). In the SOEP, individuals under 16 years do not complete their own questionnaire but are only described by the household head. Our results do not change in any substantial way if we do not omit children living in the households of recent immigrants.

Changes in household types

To establish a counterfactual income distribution in which everything is as in 2015/16 (period 1) but the distribution of household types is as in 2005/06 (period 0), we replace in the situation of 2015/16 the population shares of the different household types by those of 2005/06. Formally,

$$f_{cf}(y) = \sum_{j=1}^{6} w_{0j} f_{1j}(y), \tag{A.6}$$

where w_{0j} denote the population shares of household types $j = 1, \ldots, 6$ in period 0, and f_{1j} the distribution of net equivalent incomes of individuals living in household type $j = 1, \ldots, 6$ in period 1.

Changes in individual and household characteristics

In a similar fashion, we construct an income distribution that would have prevailed in 2015/16 if the joint distribution of individual and household characteristics x had still been as in 2005/06. To this end, we compute, separately by household type j,

$$f_{cf,j}(y) = \int_x f_{1j}(y|x) \left[\frac{dF_{0j}(x)}{dF_{1j}(x)}\right] dF_{1j}(x), \tag{A.7}$$

with reweighting factors

$$\frac{dF_{0j}(x)}{dF_{1j}(x)} = \frac{P_j(x|t=0)}{P_j(x|t=1)} = \frac{P_j(t=0|x)}{P_j(t=1|x)} \cdot \frac{P_j(t=1)}{P_j(t=0)} \tag{A.8}$$

obtained from predictions based on logit models $P_j(t=1|x)$, $P_j(t=0|x)$. We include into the logit models all the individual and household characteristics listed in table 2.2.[21] The reweighting factors are computed by household type and we include for each

[21] The estimation results for these logit models are reported in appendix A.7, see tables A.7 and A.8.

household type only the characteristics that are present in the respective type (e.g., we do not include information on children in household types without children). The terms $P_j(t = 1), P_j(1 = 0)$ are the weighted sample fractions of period 1 and 0, respectively, in the combined sample of periods 1 and 0. The final counterfactual distribution is obtained by aggregating across all household types,

$$f_{cf}(y) = \sum_{j=1}^{6} w_{1j} f_{cf,j}(y). \tag{A.9}$$

Changes in the level and structure of pay

In order to assess the effects of changes in the level and structure of the returns to labour market characteristics, we estimate wage regressions for the labour market situations in 2005/06 (period 0) and 2015/16 (period 1), separately by gender and the three employment categories (full time, part time, marginal part time).[22]

The regressions for monthly wages take the form

$$log(wage) = z\beta + u, \tag{A.10}$$

where the vector of individual characteristics z includes nationality, East German residence, disability status, age, age squared, educational qualification in three categories, work experience and work experience squared. We include in our regressions only individuals aged between 18 and 64 years who are not in education. Our counterfactual wage computations capture three aspects: i) general wage gains (reflected in the changing regression intercepts), ii) changes in wage differentials (reflected in the changes of the estimated regression coefficients $\hat{\beta}$), and iii) changes in the dispersion of unobserved (i.e., residual) wage components u.

More concretely, we carry out the following calculations:

$$\widehat{wage}_1(z, rank_1) = exp(z\hat{\beta}_1 + \hat{u}_1(rank_1)) \tag{A.11}$$

$$\widehat{wage}_0(z, rank_1) = exp(z\hat{\beta}_0 + \hat{u}_0(rank_1)) \tag{A.12}$$

[22]The detailed regression results are available in appendix A.7, see tables A.9 to A.11.

$$wage_1^{cf} = wage_1 + (\widehat{wage_0}(z, rank_1) - \widehat{wage}_1(z, rank_1)). \tag{A.13}$$

As evident from the last line, the factual wages in 2015/16 are corrected upwards/downwards by a correction term reflecting how much more/less a person with characteristics z and $rank_1$ in the residual wage distribution of period 1 would have earned in period 1 if the pay structure in period 1 had still been as in period 0. Note that all wage changes are in real terms (all wage information is expressed in prices of 2016, except for tax calculations for which we termporarily convert incomes back to nominal values).

We then multiply the counterfactual wages of each person by the actual number of months worked in the respective employment category to obtain the counterfactual annual labour income under the assumptions that the level and structure of pay in 2015/16 had been as in 2005/06. Summing within households and recomputing taxes and social security contributions yields the counterfactual annual household net income

$$\begin{aligned} y_{cf} = {} & y_{Market} + \hat{\Delta} y_{Labour} + y_{Pens} + y_{Trans} \\ & - ssc_1(y_{Labour} + \hat{\Delta} y_{Labour}, y_{Pens}) - tax_1(y_{Tax} + \hat{\Delta} y_{Labour}). \end{aligned} \tag{A.14}$$

Changes in capital incomes

For the computation of counterfactual capital incomes, we first determine the rank of the household in period 1 (2015/16) in the distribution of rental incomes, $RentRank_1$, and the rank in the distribution of other capital incomes, $CapRank_1$. We then compute the ratio of the percentiles belonging to this rank in the two distributions of period 0 and period 1 to rescale the observed value of rental and other capital incomes of period 1. This leads to the correction terms

$$\hat{\Delta} y_{Rent} = PercRent_0(RentRank_1) \frac{y_{Rent}}{PercRent_1(RentRank_1)} - y_{Rent} \tag{A.15}$$

$$\hat{\Delta} y_{Cap} = PercCap_0(CapRank_1) \frac{y_{Cap}}{PercCap_1(CapRank_1)} - y_{Cap}, \tag{A.16}$$

which we use to correct household capital incomes in order to arrive at counterfactual household net income reflecting the level and structure of capital incomes of period 0

(2005/06)

$$\begin{aligned} y_{cf} &= y_{Market} + \hat{\Delta} y_{Rent} + \hat{\Delta} y_{Cap} + y_{Pens} + y_{Trans} \\ &\quad - ssc_1(y_{Labour}, y_{Pens}) - tax_1(y_{Tax} + \hat{\Delta} y_{Rent} + \hat{\Delta} y_{Cap}). \end{aligned} \tag{A.17}$$

Changes in the transfer system

For our counterfactual simulations, we reverse the reforms listed in the main text (extension of mothers' pension, the abolition of the temporary supplement to ALG II after receipt of ALG I, the changes in child/student allowances and in unemployment benefit II) by changing the respective income component at the level of the individual in period 1 (2015/16) and by aggregating at the household level. This yields our counterfactual annual household income

$$\begin{aligned} y_{cf} &= y_{Market} + y_{Pens} + \hat{\Delta} y_{Pens} + y_{Trans} + \hat{\Delta} y_{Trans} \\ &\quad - ssc_1(y_{Labour}, y_{Pens} + \hat{\Delta} y_{Pens}) - tax_1(y_{Tax} + \hat{\Delta} y_{Pens}). \end{aligned} \tag{A.18}$$

Changes in taxes and social security contributions

To assess the effects of changes in the tax and social security system, we replace in the calculations for period 1 (2015/16) the tax and social security contributions system with that of period 0 (2005/06):

$$y_{cf} = y_{Market} + y_{Pens} + y_{Trans} - ssc_0(y_{Labour}, y_{Pens}) - tax_0(y_{Tax}) \tag{A.19}$$

For more details about our simulation of taxes and social security contributions, see appendix section A.5.

A.4 Limitations of our methodology

As pointed out in the main text, our counterfactual calculations ignore behavioural reactions and equilibrium effects. Note that such effects have often been found to be small, see Jessen (2019). Modelling such effects would necessarily rely on a large number of potentially controversial and often arbitrary assumptions. This represents a trade-off. Ignoring equilibrium effects certainly also presents a limitation, but the effects calculated by us present transparent counterfactual operations allowing us to assess the quantitative importance of different channels of distributional change irrespective of whether we attach a causal interpretation to them.

On a related note, we point out that the validity of our results is generally *not* affected by the presence of endogenous explanatory variables in our models for employment and wages. The reason is that our task is counterfactual *prediction* rather than causal modelling. The only assumption we have to maintain is that the degree of endogeneity of our regressors does not change substantially between periods 0 and 1. For example, if the correlation between education and unobservables in 2015/16 is the same as in 2005/06, we can realistically predict the wage of a person with a certain level of education in 2015/16 using the counterfactual wage schedule of 2005/06 because regression coefficients incorporate the effect of correlated unobserved components (such as ability) whose correlation with observables is, by assumption, constant over time. Any unproxied selectivity with respect to unobservables or changes therein certainly remain a limitation of the analysis. Note, however, that correcting for unobserved selectivity generally requires valid instrumental variables which are typically unavailable.

A.5 Income tax and social security contributions

Our income tax calcuations comprise the following elements:

- Joint taxation of married couples living in the same household
- Deduction of various tax allowances (*Sonderausgabenpauschale, Werbungskostenpauschale, Altersentlastungsbetrag*, contributions to pension and social security system, extra allowances for lone parents)
- Exact computation of income tax burden using tax formula of given year
- Taxes on old age pensions incl. allowances (increasing tax rate across years, *Versorgungsfreibetrag, Versorgungshöchstbetrag, Altersentlastungsbetrag*)
- Progression clause for unemployment benefit I and maternity benefits (*Progressionsvorbehalt*, i.e., these income sources are not taxed but they are added when determining the marginal tax rate)
- Child allowance: households either receive the child allowance as a direct payment,

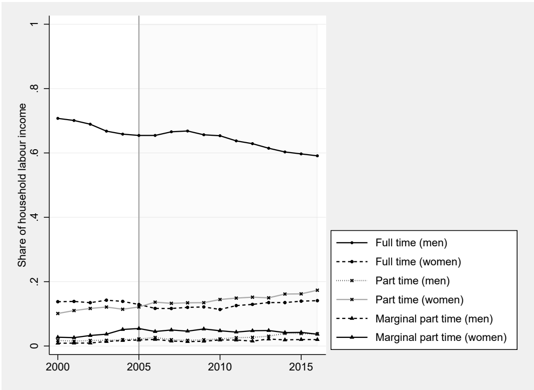
or, if more favourable, deduct child allowances from their taxable income in order to reduce their tax burden *(Günstigerprüfung Kindergeld)*

- Withholding tax on income from interest, dividends and similar income sources introduced 1st January 2009 (*Abgeltungssteuer*, flat rate of 25 percent)
- Solidarity surcharge (5.5 percent on income tax burden)

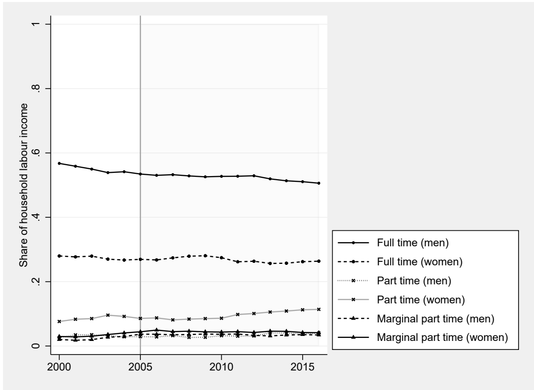
The calculation of contributions to the social security system include the following elements:

- The exact value of the social security contribution ceiling in the pension, unemployment, health and old age care insurance in each year
- The exact contribution rates in the pension, unemployment, health and old age care insurance in each year (only contributions by employees, not by employers)
- The exact contribution rates in the health and old age care insurance in each year for the income sources of pensioners that are subject to social security contributions

A.6 Additional figures

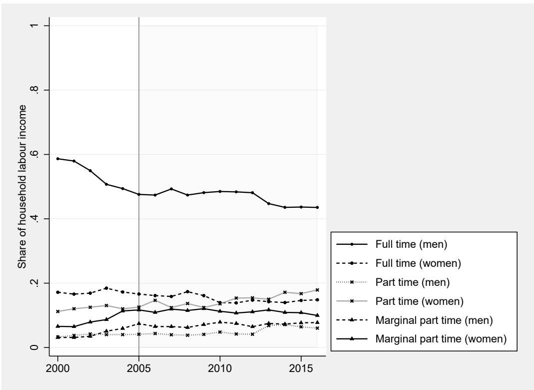


(a) With children

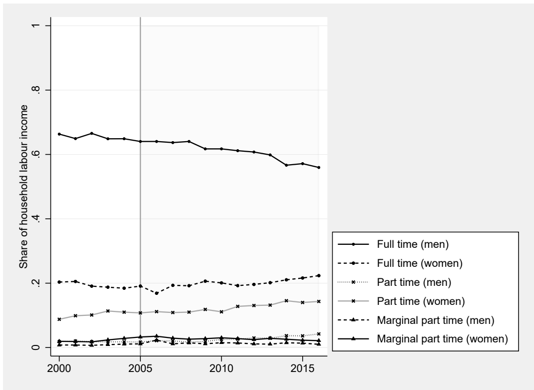


(b) Without children

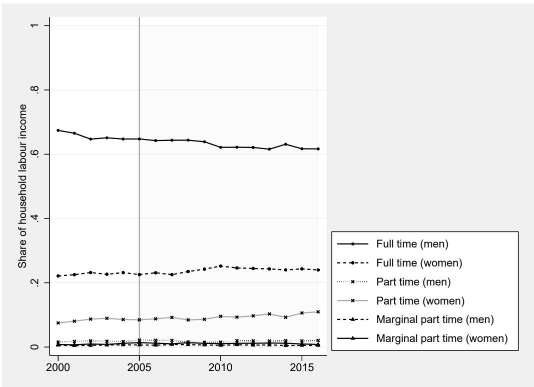
Figure A.1 – Share of household labour income contributed by men/women
Shares of household labour income contributed by men/women from different employment categories in households with and without children (Source: Socio-Economic Panel, own calculations)



(a) Lower tercile



(b) Middle tercile



(c) Upper tercile

Figure A.2 – Share of household labour income contributed by men/women by tercile
Shares of household labour income contributed by men/women from different employment categories in the lower, middle and upper part of the distribution of net equivalised incomes (Source: Socio-Economic Panel, own calculations)

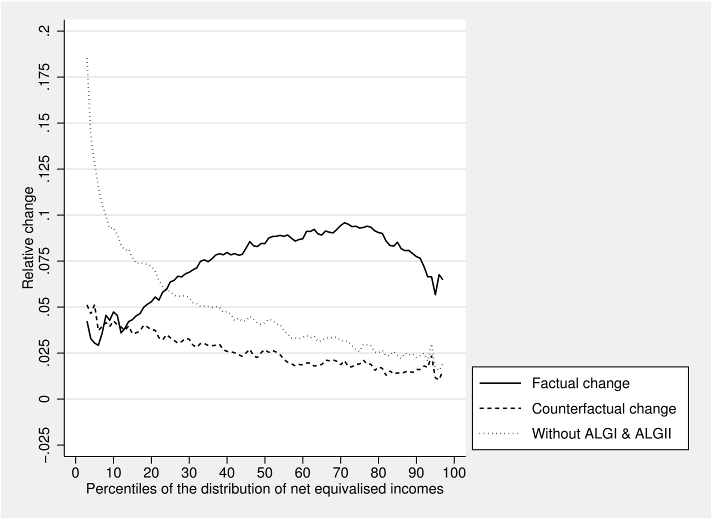


Figure A.3 – Robustness (i): alternative definition of employment intensity
Robustness analysis (i): relative change of percentiles of net equivalised income due to the employment boom, logit specification (zero = 0 to 5 months worked, one = 6 to 12 months worked) and rounding correction to obtain intermediate values 1 to 11 months for predicted months worked
(Source: Socio-Economic Panel, own calculations)

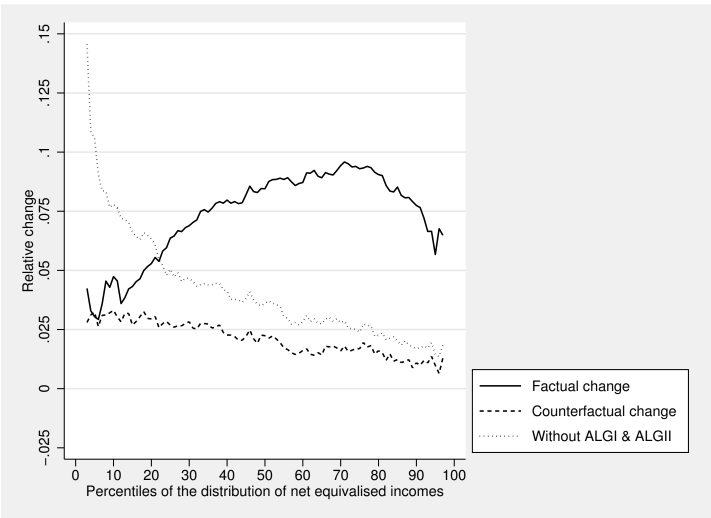


Figure A.4 – Robustness (ii): specification without rounding correction
Robustness analysis (ii): relative change of percentiles of net equivalised income due to the employment boom, logit specification as in paper (zero = 0 months worked, one = 1 to 12 months worked) but no rounding correction to obtain intermediate values 1 to 11 months for predicted months worked (i.e., counterfactual predictions are either zero or twelve months) (Source: Socio-Economic Panel, own calculations)

A.7 Additional tables

	Men		Women	
Variable	2005/06	2015/16	2005/06	2015/16
Foreign nationality	-0.613	0.133	-0.242	0.067
	[0.184]	[0.172]	[0.171]	[0.134]
East Germany	-0.338	-0.242	0.390	0.277
	[0.117]	[0.143]	[0.112]	[0.128]
Disability	-1.255	-1.247	-0.719	-0.904
	[0.132]	[0.158]	[0.161]	[0.137]
Age	0.208	0.272	0.061	0.050
	[0.043]	[0.040]	[0.035]	[0.033]
Age^2	-0.004	-0.004	-0.002	-0.001
	[0.000]	[0.000]	[0.000]	[0.000]
University degree	1.570	0.657	0.812	0.450
	[0.201]	[0.191]	[0.158]	[0.148]
Abitur and/or voc. train.	0.797	0.393	0.161	-0.012
	[0.153]	[0.153]	[0.135]	[0.132]
Years of work experience	0.150	0.067	0.107	0.045
	[0.048]	[0.037]	[0.025]	[0.027]
Years of work experience^2	-0.001	0.000	-0.003	-0.001
	[0.002]	[0.001]	[0.001]	[0.001]
FT months in past 3 years	0.094	0.122	0.111	0.138
	[0.003]	[0.005]	[0.004]	[0.004]
PT months in past 3 years	-0.081	-0.045	-0.034	-0.017
	[0.010]	[0.008]	[0.005]	[0.005]
MPT months in past 3 years	-0.043	-0.031	-0.037	-0.017
	[0.008]	[0.008]	[0.006]	[0.005]
UNEM months in past 3 years	-0.066	-0.045	-0.035	-0.026
	[0.006]	[0.008]	[0.009]	[0.006]
Fraction of HH age < 3	-0.178	-0.294	-2.799	-2.326
	[0.165]	[0.234]	[0.227]	[0.265]
Fraction of HH $3 <$ age ≤ 6	0.078	-0.103	-0.729	-0.597
	[0.173]	[0.176]	[0.196]	[0.127]
Fraction of HH $6 <$ age ≤ 17	-0.146	-0.182	-0.509	-0.427
	[0.082]	[0.066]	[0.069]	[0.061]
Constant	-2.834	-4.754	-1.571	-1.370
	[0.733]	[0.706]	[0.688]	[0.617]
N	12,751	15,267	13,936	18,117
No. cluster	6,876	8,811	7,636	10,179

Standard errors in parenthesis (clustered at the household level)

Table A.4 – Logit estimation results (full-time employment)
Logit models for full-time employment, 0=zero months per year,
1=one to twelve months per year (Source: Socio-Economic Panel, own calculations)

	Men		Women	
Variable	2005/06	2015/16	2005/06	2015/16
Foreign nationality	-0.267	0.444	-0.331	-0.032
	[0.344]	[0.234]	[0.192]	[0.169]
East Germany	-0.319	-0.008	0.013	0.099
	[0.192]	[0.209]	[0.146]	[0.128]
Disability	-0.725	-1.159	-1.216	-1.283
	[0.283]	[0.307]	[0.222]	[0.223]
Age	0.027	0.096	0.135	0.198
	[0.060]	[0.060]	[0.044]	[0.039]
Age^2	-0.002	-0.002	-0.003	-0.003
	[0.001]	[0.001]	[0.000]	[0.000]
University degree	1.519	0.988	0.871	0.951
	[0.322]	[0.284]	[0.192]	[0.179]
Abitur and/or voc. train.	0.463	0.211	0.512	0.336
	[0.264]	[0.230]	[0.144]	[0.158]
Years of work experience	0.130	-0.035	0.181	0.066
	[0.065]	[0.050]	[0.025]	[0.025]
Years of work experience^2	-0.001	0.002	-0.003	-0.001
	[0.002]	[0.002]	[0.001]	[0.001]
FT months in past 3 years	0.035	0.086	0.068	0.063
	[0.009]	[0.009]	[0.005]	[0.006]
PT months in past 3 years	0.126	0.175	0.116	0.152
	[0.011]	[0.011]	[0.005]	[0.006]
MPT months in past 3 years	0.008	0.044	-0.037	0.002
	[0.012]	[0.012]	[0.007]	[0.006]
UNEM months in past 3 years	-0.039	-0.008	-0.038	-0.006
	[0.011]	[0.010]	[0.008]	[0.006]
Fraction of HH age < 3	0.389	0.022	-0.851	-0.780
	[0.243]	[0.200]	[0.145]	[0.154]
Fraction of HH $3 < \text{age} \leq 6$	0.085	0.001	0.365	0.163
	[0.248]	[0.205]	[0.128]	[0.108]
Fraction of HH $6 < \text{age} \leq 17$	-0.013	0.128	0.054	-0.215
	[0.114]	[0.091]	[0.075]	[0.068]
Months in FT	-0.422	-0.447	-0.620	-0.602
	[0.024]	[0.026]	[0.030]	[0.029]
Constant	-1.663	-4.190	-4.233	-4.891
	[1.074]	[1.041]	[0.875]	[0.777]
N	12,751	15,267	13,936	18,117
No. cluster	6,876	8,811	7,636	10,179

Standard errors in parenthesis (clustered at the household level)

Table A.5 – Logit estimation results (part-time employment)
Logit models for part-time employment, 0=zero months per year,
1=one to twelve months per year (Source: Socio-Economic Panel, own calculations)

	Men		Women	
Variable	2005/06	2015/16	2005/06	2015/16
Foreign nationality	-0.507	-0.127	-0.100	0.131
	[0.282]	[0.205]	[0.183]	[0.146]
East Germany	-0.524	-0.230	-0.150	-0.299
	[0.145]	[0.148]	[0.129]	[0.135]
Disability	-0.539	-0.233	-0.701	-0.747
	[0.227]	[0.204]	[0.202]	[0.169]
Age	-0.004	0.010	-0.008	0.000
	[0.063]	[0.056]	[0.039]	[0.034]
Age^2	-0.001	-0.000	-0.001	-0.000
	[0.001]	[0.001]	[0.000]	[0.000]
University degree	0.577	-0.066	0.242	0.089
	[0.234]	[0.210]	[0.186]	[0.162]
Abitur and/or voc. train.	0.141	-0.372	0.182	0.391
	[0.189]	[0.175]	[0.143]	[0.127]
Years of work experience	0.155	0.056	0.101	0.055
	[0.065]	[0.055]	[0.028]	[0.023]
Years of work experience^2	-0.002	-0.002	-0.002	-0.001
	[0.002]	[0.002]	[0.001]	[0.001]
FT months in past 3 years	0.011	0.046	0.021	0.039
	[0.007]	[0.008]	[0.006]	[0.006]
PT months in past 3 years	0.060	0.060	0.045	0.058
	[0.012]	[0.012]	[0.006]	[0.006]
MPT months in past 3 years	0.140	0.145	0.134	0.133
	[0.008]	[0.007]	[0.006]	[0.005]
UNEM months in past 3 years	0.000	-0.002	-0.006	0.003
	[0.007]	[0.008]	[0.006]	[0.005]
Fraction of HH age < 3	0.035	0.335	-0.824	-0.071
	[0.193]	[0.178]	[0.150]	[0.147]
Fraction of HH $3 <$ age ≤ 6	0.135	-0.107	0.354	0.104
	[0.236]	[0.182]	[0.139]	[0.122]
Fraction of HH $6 <$ age ≤ 17	-0.023	0.130	-0.096	-0.045
	[0.093]	[0.066]	[0.073]	[0.063]
Months in FT	-0.230	-0.265	-0.271	-0.290
	[0.019]	[0.019]	[0.018]	[0.016]
Months in PT	-0.291	-0.324	-0.310	-0.285
	[0.038]	[0.039]	[0.020]	[0.015]
Constant	-1.288	-2.209	-0.688	-1.488
	[1.080]	[0.984]	[0.734]	[0.648]
N	12,751	15,267	13,936	18,117
No. cluster	6,876	8,811	7,636	10,179

Standard errors in parenthesis (clustered at the household level)

Table A.6 – Logit estimation results (marginal part-time employment)
Logit models for marginal part-time employment, 0=zero months per year, 1=one to twelve months per year (Source: Socio-Economic Panel, own calculations)

HH-type	1	2	3
Female head of HH	-0.101		
	[0.100]		
Foreign nationality head of HH	1.008		0.852
	[0.331]		[0.153]
University degree head of HH	0.994	0.599	
	[0.145]	[0.145]	
Abitur and/or voc. training head of HH	0.570	0.174	
	[0.101]	[0.130]	
Years of work experience head of HH	0.043	0.025	0.019
	[0.006]	[0.008]	[0.005]
Age head of HH	-0.057	0.053	
	[0.011]	[0.008]	
Age head of HH^2	0.001		
	[0.000]		
East Germany head of HH		-0.396	
		[0.104]	
Spouse/2nd person female		-0.908	
		[0.119]	
Spouse/2nd person foreign		0.752	
		[0.246]	
Spouse/2nd person Abitur and/or voc. training		0.072	
		[0.089]	
Spouse/2nd person years of work experience		0.018	
		[0.004]	
Constant	-0.019	-4.024	-0.150
	[0.276]	[0.645]	[0.068]
N	4,964	13,637	8,092
No. cluster	2,430	3,370	4,582

Standard errors in parenthesis (clustered at the household level)

Table A.7 – Logit estimation results for reweighting, households without children
Logit models for reweighting (0=in 2005/06, 1=in 2015/16), household types: (1) single pensioner households, (2) multiple pensioner households and (3) single adult households without children
(Source: Socio-Economic Panel, own calculations)

HH-type	4	5	6
Female head of HH	0.240	-0.907	
	[0.116]	[0.280]	
Foreign nationality head of HH	0.446		
	[0.151]		
University degree head of HH	0.446		0.363
	[0.123]		[0.085]
Abitur and/or voc. training head of HH	0.297		
	[0.110]		
Years of work experience head of HH	0.022	-0.079	-0.023
	[0.007]	[0.044]	[0.026]
Years of work experience head of HH^2		0.005	0.002
		[0.002]	[0.001]
Age head of HH	-0.048		
	[0.013]		
Age head of HH^2	0.000		
	[0.000]		
East Germany head of HH	-0.201	0.232	
	[0.076]	[0.182]	
Spouse/2nd person female	-0.210		
	[0.119]		
Spouse/2nd person foreign	-0.533		-0.135
	[0.154]		[0.107]
Spouse/2nd person university degree	0.308		-0.236
	[0.102]		[0.125]
Spouse/2nd person Abitur and/or voc. training			-0.469
			[0.098]
Spouse/2nd person years of work experience	-0.088		-0.058
	[0.012]		[0.021]
Spouse/2nd person years of work experience^2	0.003		0.003
	[0.000]		[0.001]
Spouse/2nd person age	0.034		-0.040
	[0.013]		[0.033]
Spouse/2nd person age^2	-0.000		0.001
	[0.000]		[0.000]
Fraction of HH age < 3			0.295
			[0.080]
Fraction of HH $3 < \text{age} \leq 6$			0.186
			[0.079]
$>$ 2 adults in HH			-0.231
			[0.100]
Constant		0.799	0.065
		[0.324]	[0.598]
N	27,782	6,508	55,801
No. cluster	6,948	1,598	7,965

Standard errors in parenthesis (clustered at the household level)

Table A.8 – Logit estimation results for reweighting, households with children

Logit models for reweighting, household types: (4) multiple adult household without children, (5) single adult household with children and (6) multiple adult household with children (Source: Socio-Economic Panel, own calculations)

	Men		Women	
Variable	2005/06	2015/16	2005/06	2015/16
Foreign nationality	-0.044	-0.159	-0.119	-0.148
	[0.034]	[0.026]	[0.048]	[0.037]
East Germany	-0.404	-0.379	-0.360	-0.237
	[0.021]	[0.024]	[0.031]	[0.027]
Disability	-0.077	-0.118	-0.038	-0.063
	[0.034]	[0.035]	[0.058]	[0.042]
Age	0.035	0.065	0.036	0.019
	[0.010]	[0.009]	[0.012]	0.009
Age^2	-0.000	-0.001	-0.000	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]
University degree	0.619	0.628	0.706	0.669
	[0.036]	[0.033]	[0.053]	[0.044]
Abitur and/or voc. train.	0.165	0.174	0.244	0.286
	[0.029]	[0.028]	[0.050]	[0.039]
Years of work experience	0.038	0.020	0.024	0.022
	[0.010]	[0.008]	[0.010]	[0.008]
Years of work experience^2	-0.001	-0.001	-0.001	-0.001
	[0.000]	[0.000]	[0.000]	[0.000]
Constant	6.876	6.311	6.698	6.932
	[0.181]	[0.184]	[0.223]	[0.174]
N	10,265	11,278	5,047	5,815
No. cluster	5,726	6,336	3,085	3,541

Standard errors in parenthesis (clustered at the household level)

Table A.9 – Full-time wage regressions
Estimation results for monthly log full-time wages
(Source: Socio-Economic Panel, own calculations)

	Men		Women	
Variable	2005/06	2015/16	2005/06	2015/16
Foreign nationality	-0.042	-0.025	-0.130	-0.126
	[0.133]	[0.092]	[0.063]	[0.063]
East Germany	-0.193	-0.042	-0.053	-0.046
	[0.103]	[0.082]	[0.038]	[0.045]
Disability	-0.129	-0.023	0.156	0.052
	[0.123]	[0.099]	[0.094]	[0.059]
Age	0.049	0.003	0.019	0.036
	[0.031]	[0.029]	[0.016]	[0.012]
Age^2	-0.001	-0.000	-0.000	-0.001
	[0.000]	[0.000]	[0.000]	[0.000]
University degree	0.319	0.673	0.672	0.542
	[0.143]	[0.120]	[0.058]	[0.048]
Abitur and/or voc. train.	0.117	0.207	0.166	0.164
	[0.124]	[0.095]	[0.047]	[0.042]
Years of work experience	0.024	0.069	0.040	0.032
	[0.030]	[0.025]	[0.008]	[0.006]
Years of work experience^2	0.000	-0.002	-0.001	-0.000
	[0.001]	[0.001]	[0.000]	[0.000]
Constant	5.935	6.789	6.542	6.287
	[0.547]	[0.511]	[0.319]	[0.235]
N	516	887	3,456	5,797
No. cluster	407	667	2,243	3,562

Table A.10 – Part-time wage regressions
Estimation results for monthly log part-time wages
(Source: Socio-Economic Panel, own calculations)

	Men		Women	
Variable	2005/06	2015/16	2005/06	2015/16
Foreign nationality	0.347	0.271	-0.006	0.081
	[0.125]	[0.080]	[0.063]	[0.060]
East Germany	-0.178	-0.152	-0.379	-0.193
	[0.091]	[0.102]	[0.064]	[0.076]
Disability	-0.196	0.234	-0.280	-0.075
	[0.131]	[0.114]	[0.097]	[0.095]
Age	-0.051	-0.040	-0.018	-0.006
	[0.025]	[0.025]	[0.016]	[0.014]
Age^2	0.000	0.000	0.000	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]
University degree	-0.038	-0.076	0.083	-0.214
	[0.145]	[0.141]	[0.096]	[0.083]
Abitur and/or voc. train.	-0.013	0.018	0.068	-0.015
	[0.101]	[0.096]	[0.054]	[0.053]
Years of work experience	0.005	0.040	0.020	0.023
	[0.030]	[0.026]	[0.012]	[0.009]
Years of work experience^2	0.001	-0.001	-0.000	-0.000
	[0.001]	[0.001]	[0.000]	[0.000]
Constant	6.730	6.141	6.169	5.851
	[0.430]	[0.426]	[0.295]	[0.277]
N	801	1,251	2,149	3,015
No. cluster	630	972	1,472	2,092

Table A.11 – Marginal part-time wage regressions
Estimation results for monthly log marginal part-time wages
(Source: Socio-Economic Panel, own calculations)

CHAPTER 3

Selectivity-corrected wage distributions and the evolution of the German gender wage gap*

3.1 Introduction

Since the seminal works of Heckman (1974, 1979), the issue of unobserved selectivity has been a central concern in labour economics. It is well-established that selection on unobservables can bias the measurement of wage disparities between groups of workers if one group is more positively or negatively selected than the other. For instance, the wage gap between men and women will be underestimated if women with less favourable unobserved characteristics are more likely to stay out of the labour market. A growing body of literature has examined the effects of such unobserved selectivity on the gender

*This chapter is based on: Biewen, Martin, Pascal Erhardt, Bernd Fitzenberger and Miriam Sturm (2025): Selectivity-corrected wage distributions and the evolution of the German gender wage gap, *unpublished manuscript, University of Tübingen.* We thank Blaise Melly, Peter Haan and the audiences at the World SOLE-EALE-AASLE Conference 2025 and the Annual Meeting of the Verein für Socialpolitik 2025 for helpful discussions. Nicolas Erhardt, Maximiliane Sick and Lara Müller provided excellent research assistance.

gap in full-time wages (e.g., Albrecht et al., 2009; Arellano and Bonhomme, 2017; Mulligan and Rubinstein, 2008; Olivetti and Petrongolo, 2008, see literature review below). For a long time, selection correction was targeted at women, whose self-selection behaviour and the resulting labour market participation – which is traditionally lower than that of men – was typically assumed to be the main driver of the bias. This view has shifted in light of the fluctuations in male employment shares in many industrialised countries caused by events such as the global financial crisis (Dolado et al., 2020; Elass, 2024; Maasoumi and Wang, 2019). A distinguishing feature of recent contributions is that the methodological advances of Arellano and Bonhomme (2017) and Chernozhukov et al. (2025) allow researchers to investigate the impact of unobserved selectivity in a manner that extends beyond mean or median outcomes, encompassing entire outcome distributions.

This paper contributes to the expanding literature by providing new evidence on the role of unobserved selectivity in shaping male and female wage distributions in Germany. Specifically, we contribute threefold. Firstly, in contrast to the majority of preceding studies on this topic, the present paper appears to be the first empirical application of a distribution regression model with sample selection correction, as proposed by Chernozhukov et al. (2025). By choosing this method, we are able to flexibly investigate the evolution of unobserved selectivity patterns across the entire outcome distribution. Secondly, our analysis is based on high-quality German administrative data that have been used in a number of influential papers on the wage distribution, including those by Dustmann et al. (2009), Card et al. (2013), Dustmann et al. (2022), and Bossler and Schank (2023). One challenge in using administrative data to study unobserved selectivity is the lack of household variables that can be used as instruments. We propose using group variables based on labour market dynamics as instruments for unobserved selectivity. In addition, our study spans distinct phases of the business cycle (recession 2000-2005, labour market boom 2012-2017), enabling us to examine the influence of the cycle on unobserved selectivity. Thirdly, in contrast to the majority of previous studies, we consider unobserved selectivity in both full-time and part-time wages, whereas the existing literature has almost exclusively focussed on the full-time case. Given the substantial role of female part-time employment in many industrialised countries and the growing importance of male part-time employment, the part-time case appears increasingly relevant.

Our results suggest an important role for unobserved selectivity in male and female wages in Germany. For full-time men, selection on unobservables was largely neutral over a wide range of the distribution in the past, but turned negative as more men were drawn into a booming labour market after 2012. In contrast, male selectivity tends to be positive at the lower end of the wage distribution, potentially due to the protective effect of Germany's generous social safety net. For full-time women, we find generally negative unobserved selectivity, which may be attributed to assortative matching in combination with the German tax and transfer system that discourages secondary earnings. Again, selectivity is less negative towards the bottom of the distribution. In the case of part-time work, our results show negative selectivity for men, while for women, selectivity patterns are more complex, shifting from negative in the lower half to positive in the upper part of the distribution. Our findings reveal pronounced heterogeneity in selectivity patterns across the wage distribution, which has not been recognised by the previous literature. Gender gaps in full-time as well as in part-time wages have narrowed between 2000-2005 and 2012-2017. For the full-time gender wage gap, declining differences in unobserved selectivity between men and women and improved observables for women explain most of the change. For part-time work, declining differences in unobserved selectivity also play a significant role, but most of the changes are explained by a convergence in wage returns to male and female part-time jobs.

3.2 Related literature

A growing body of literature has studied wage differences between men and women accounting for unobserved selectivity. There is a broad consensus in the literature that conventional measures of the gender pay gap may over- or underestimate inequality between men and women if such selectivity is not considered. Researchers have proposed different approaches to recover wage distributions that are purged of selection. Important early contributions focussing on differences in mean or median wages include Mulligan and Rubinstein (2008) and Olivetti and Petrongolo (2008). Employing selectivity corrections based on Heckman (1979), Mulligan and Rubinstein (2008) find that female selectivity in full-time work in the US shifted from negative to positive in the 1990s, contributing to the narrowing of the observed gender gap (see Beblo et al., 2003, for European countries). Olivetti and Petrongolo (2008) impute the unobserved wages of the non-employed to study the impact on selection on wage gaps in OECD countries. Their findings indicate limited effects in most countries, with the exception of southern Europe, where positive selection of women reduced observed gender gaps. A recent study following a similar methodology is Blau et al. (2024). They conclude that correcting for unobserved selection leads to larger declines of the US gender wage gap over time than without correction, implying that changes in selectivity played a role in narrowing the gap.

Extending the analysis beyond the median gender gap, Albrecht et al. (2009) and Chzhen and Mumford (2011) study the impact of unobserved selectivity across the entire wage distribution. In an application for the Netherlands, Albrecht et al. (2009) find positive selection of women into full-time work, reducing observed wage differences between men and women. Chzhen and Mumford (2011) obtain similar results for the UK. Both studies apply a selectivity correction for quantile regression models developed by Buchinsky (1998, 2001). This correction method was later shown to be highly restrictive by Huber and Melly (2015). Addressing the problem pointed out by this analysis, Biewen et al. (2020) propose a modification to the method of Albrecht et al. (2009). Their application to the full-time gender wage gap in Germany still produces similar findings as in Albrecht et al. (2009) and Chzhen and Mumford (2011). Employing a related method, Fitzenberger and de Lazzer (2022) examine the effect of unobserved selectivity on male full-time wages in Germany. Their results suggest positive selection of men into full-time work, which became less positive as more men entered the expanding labour market.

In a significant methodological advance, Arellano and Bonhomme (2017) developed a direct method for correcting entire distributions of outcomes for unobserved selectivity. Their approach entails the selection specific rotation of the quantile indices of individual observations. This is achieved by modelling the joint distribution of unobservables in participation and wage equation by a bivariate copula. Arellano and Bonhomme (2017) apply their quantile regression method to male and female wages in the UK. In contrast to previous contributions, they also explicitly model unobserved selectivity for males, finding positive selection on unobservables for both men and women. Maasoumi and Wang (2019) apply Arellano and Bonhomme's (2017) method to US data. Their results point to negative male and female selection into full-time work that turned positive in the 1990s. Chen et al. (2024) propose a modification of Arellano and Bonhomme's (2017) original method. They obtain significant negative selection among males and positive selection for females based on the same data as Arellano and Bonhomme (2017). Arellano and Bonhomme's (2017) method has been applied to data from various countries, leading to diverse findings. Dolado et al. (2020) use the method to examine employment and wage patterns in EU countries before and after the Great Recession. The authors obtain varied results for different countries, suggesting that male selection became positive during the recession, while positive female selection decreased in some countries due to an added-worker effect (low-ability women joined the labour market in the recession). Also using the method developed by Arellano and Bonhomme (2017), Elass (2024) presents results for France, Finland, and the UK: her analysis implies positive male and female selection on unobservables for France, whereas it is negative for Finland and the UK. Pereda-Fernández (2025) applies Arellano and Bonhomme's (2017) method to the US, finding positive selection for both men and women, but a differential evolution of male and female selectivity patterns reduces the observed gender wage gap over time.

The quantile regression approach of Arellano and Bonhomme (2017) imposes the restriction of a constant degree of selectivity over all quantile indices, as represented by the scalar copula parameter. Recently, Chernozhukov et al. (2025) have proposed an alternative model based on a bivariate probit that allows for heterogeneity in the selection structure to be captured. In section 3.3, we give a more detailed description of this method. In their application to the UK, Chernozhukov et al. (2025) find significant differences in selectivity across the distribution: in the past, male unobserved selectivity was negative at the bottom and positive at the top, but became more uniformly positive in more recent years. In contrast, female selectivity is generally negative, although it has recently become less pronounced and more uniform. These results indicate that methods assuming homogeneous selectivity may miss important aspects of the selection behaviours of men and women. The present paper contributes to this recent strand of the literature by employing Chernozhukov et al.'s (2025) method to study selectivity patterns in male and female wage distributions in Germany. Apart from full-time wages, we also examine selectivity in part-time wages. To the best of our knowledge, only a limited number of previous contributions have considered the issue of unobserved selectivity with regard to part-time wages. With the exception of Gallego-Granados (2019), who employs a distributional imputation technique due to Melly and Santangelo (2014), the majority of these studies use elementary techniques for selection correction (Bardasi and

Gornick, 2008; Manning and Petrongolo, 2008; Matteazzi et al., 2014).[1]

We conclude by mentioning a related body of literature that focusses on selectivity at the intensive margin (working hours). For further details, the reader is referred to Fernandez-Val et al. (2023, 2024a, 2024b). We emphasise that our dataset only allows for a distinction between full-time and part-time, which prevents us from applying these techniques.

3.3 Econometric method

This section presents a short description of the method proposed by Chernozhukov et al. (2025) to estimate selectivity-corrected wage distributions. The method allows for more flexible patterns of selectivity compared to previous approaches (Albrecht et al., 2009; Buchinsky, 1998; Heckman, 1974) as the sign and the magnitude of selection may differ across the outcome distribution.

Consider the following selection process

$$\begin{aligned} D &= 1(D^* \leq 0), \\ Y &= Y^* \text{ if } D = 1, \end{aligned}$$

where Y^* is the wage a person would receive if she decided to work. The observability of the wage offer is contingent upon the individual being employed, which in turn depends on the latent propensity to work D^* (related to the difference between the wage offer and the reservation wage of the person). It is crucial to account for selectivity as the distribution of observed outcomes Y typically differs from the distribution of latent outcomes Y^* due to the endogeneity of employment decisions (because employment decisions D^* depend on wage offers Y^*). For example, if, in a group of potential workers, individuals with low wage offers opt out of employment, the resulting wage distribution appears more favourable than it is in reality, since lower wage levels are not represented among those who are actually employed.

Let F_{Y^*} and F_{D^*} denote the marginal cumulative distribution functions (CDFs) of Y^*

[1] Gallego-Granados and Wrohlich (2019) apply the method of Melly and Santangelo (2014) to full-time wages.

and D^*, respectively, and let F_{Y^*,D^*} be their joint CDF. Chernozhukov et al. (2025) show that the latter can generally be represented by a standard bivariate normal distribution with local correlation parameter ρ evaluated at point (y, d). Notably, this representation does not assume the joint distribution of Y^* and D^* to be normal. Moreover, it establishes that any joint distribution function of two random variables can be represented by a sequence of bivariate normals. Chernozhukov et al. (2025) call this unique representation the Local Gaussian Representation (LGR):

$$F_{Y^*,D^*}(y, d) = \Phi_2(\mu(y), \nu(d); \rho(y, d)), \tag{3.1}$$

where Φ_2 denotes the bivariate standard normal CDF, $\mu(y) = \Phi^{-1}(F_{Y^*}(y))$ and $\nu(d) = \Phi^{-1}(F_{D^*}(d))$. The parameter $\rho(y, d) \in [-1, 1]$ measures the local dependence between the two dichotomous variables $1(Y^* \leq y)$ and $1(D^* \leq d)$ at the threshold tuple (y, d). In the context of our present labour market application, the parameter $\rho(y, 0)$ quantifies the degree to which the occurrence of an individual being employed, (i.e., $1(D^* \leq 0)$) and concurrently receiving a wage offer below a specified threshold y (i.e., $1(Y^* \leq y)$) are correlated. Non-zero correlations cause a selection bias in the sense that the distribution of observed wages F_Y differs from the distribution of wage offers F_{Y^*} that would be observed in the hypothetical scenario in which every individual is employed.

Chernozhukov et al. (2025) demonstrate that point identification of the parameters in equation (3.1) is achieved using at least one binary instrumental variable Z_1. To this end, define a vector of covariates X determining wage offers Y^* and let $Z = (Z_1, X)$. We impose the exclusion restriction that instruments Z_1 influence the propensity to work, but not the wage offers, nor their correlation with the propensity to work. It follows that

$$\begin{aligned} F_{Y^*,D^*|Z}(y, d|z) &= \Phi_2(\mu(y|z), \nu(d|z); \rho(y, d|z)) \\ &= \Phi_2(\mu(y|x), \nu(d|z); \rho(y, d|x)). \end{aligned} \tag{3.2}$$

Chernozhukov et al. (2025) propose to parametrise the unknown parameters in (3.2) by $\mu(y|x) = -x'\beta(y)$, $\nu(d|z) = -z'\pi$ and $\rho(y, d|x) = \rho(x'\delta(y)) = \tanh(x'\delta(y))$ by the Fisher link. For the remainder, these will be referred to as outcome, selection, and sorting equations, respectively.[2]

The joint CDF of Y^*, D^* conditional on Z can be represented in terms of its LGR:

$$F_{Y^*,D^*}(y, 0|Z = z) = \Phi_2(-x'\beta(y), -z'\pi; \rho(x'\delta(y))) \tag{3.3}$$

$$= P(Y \leq y, D = 1|Z = z), \tag{3.4}$$

where the second line of this equation implies that the joint distribution $F_{Y^*,D^*|Z}$ can be recovered by estimating a series of selection-corrected probit models for wages being below a fine grid of thresholds in the support of the outcome $y \in \mathcal{Y}$. This representation of the selection problem motivates the use of a distribution regression approach of the form

$$F_{Y|X}(y|z) = \mathrm{E}[1(Y \leq y)|Z = z, D = 1], \tag{3.5}$$

where Z denotes covariates and instruments, and $F_{Y|Z}$ the CDF of observed wages Y conditional on Z. This representation of the selection problem is highly flexible as it allows for heterogeneous returns $\beta(y)$ as well as for heterogeneous selection sorting $\rho(x'\delta(y))$ across the wage distribution. Moreover, selection patterns may vary depending on the covariates, adding another layer of heterogeneity.

The model can be used to recover a number distributions of interest. Firstly, one may wish to recover the wage offer distribution, i.e., the distribution of wages that is free of selection bias:

$$F_{Y^*}(y) = \int \Phi(y|Z = z)dF_Z(z) = \int \Phi(-x'\beta(y))dF_X(x), \tag{3.6}$$

which can be estimated by its empirical counterpart

$$\hat{F}_{Y^*}(y) = \frac{1}{N}\sum_{i=1}^{N} \Phi(-X_i'\hat{\beta}(y)), \tag{3.7}$$

where Φ stands for the standard normal CDF. We will use this distribution below to describe differences in the distribution of male and female wages free of selection bias,

[2] Chernozhukov et al. (2025) add the minus signs to facilitate the interpretation of the respective parameters in terms of comparability with the classical Heckman selection model, where selection is determined by $D^* > 0$ rather than by $D^* \leq 0$.

and to describe changes in wage selectivity over time.[3]

The distribution of the observed outcome Y implied by the model is given by

$$F_Y(y) = \int \frac{\Phi_2(-x'\beta(y), z'\pi; -\rho(x'\delta(y))}{\Phi(z'\pi)} dF_Z(z|D=1) \tag{3.8}$$

$$= \frac{\int \Phi_2(-x'\beta(y), z'\pi; -\rho(x'\delta(y)))dF_Z(z)}{\int \Phi(z'\pi)dF_Z(z)}, \tag{3.9}$$

the sample analog of which is

$$\hat{F}_Y(y) = \frac{\sum_{i=1}^{N} \Phi_2(-X'\hat{\beta}(y), Z_i'\hat{\pi}; -\rho(X_i'\hat{\delta}(y)))}{\sum_{i=1}^{N} \Phi(Z_i'\hat{\pi})}. \tag{3.10}$$

Equation (3.10) can be used for the construction of counterfactual distributions of the observed outcome, which may serve as the basis of decomposition exercises exploring the sources of wage differences between groups of workers (men and women), or across periods. Depending on the question at hand, the desired counterfactuals can be constructed by combining the estimated $\hat{\beta}(y)$, $\hat{\pi}$ and $\rho(x'\hat{\delta}(y))$, as well as the marginal distributions of covariates, $\hat{F}_Z$, of the two groups accordingly. For example, in the context of the gender wage gap, we may ask what the distribution of female wages would look like if women sorted into employment like men, thereby counterfactually changing $\rho(x'\hat{\delta}(y))$ to that of men. Equation (3.10) also serves as a specification check for the model, as this distribution should coincide with the empirically observed outcome distribution if the model is correctly specified. This was the case in our application.

To obtain estimates for $\beta(y)$, π and $\rho(x'\delta(y))$, we employ the two-step procedure suggested by Chernozhukov et al. (2025). The first step is analogous to the classical Heckman selection model and consists of a probit model for selection into employment to estimate the parameters in the selection equation:

$$\hat{\pi} = \arg\max_{c \in \mathbb{R}^{d_\pi}} L_1(c) = \frac{1}{N} \sum_{i=1}^{N} [D_i \log\Phi(Z_i'c) + (1-D_i)\log\Phi(-Z_i'c)]. \tag{3.11}$$

The second step comprises a sequence of selection-corrected probit regressions over a

[3]In order to ensure the monotonicity of constructed distributions (3.7), we use the method of monotone rearrangement as in Chernozhukov et al. (2025).

finite grid of thresholds $y \in \mathcal{Y}$ in order to estimate $\beta(y)$ and $\rho(x'\delta(y))$

$$\hat{\theta}_y = \underset{t=(b,d)\in\Theta}{\arg\max}\ L_2(t,\hat{\pi}) = \frac{1}{N}\sum_{i=1}^{N} D_i[I_{yi}\log\Phi_2(-X_i'b, Z_i'\hat{\pi}; -\rho(X_i'd)) + (1-I_{yi})\log\Phi_2(X_i'b, Z_i'\hat{\pi}; \rho(X_i'd))], \tag{3.12}$$

where $I_{yi} = 1(Y_i \leq y)$ if $D_i = 1$ indicates for selected, i.e., employed, individuals whether their wage is below the given threshold y. Pointwise standard errors of the model parameters in all equations can be computed based on the usual asymptotic expansions. The estimated standard errors of $\beta(y)$ and $\rho(x'\delta(y))$ need to be adjusted for the use of the first-step estimates. Uniform confidence bands for functionals of the parameters are obtained by the multiplier bootstrap described in Chernozhukov et al. (2013, 2025). In appendix B.1, we provide detailed steps of the algorithm.

3.4 Data

Our analysis is based on administrative data from the Sample of Integrated Labour Market Biographies version 1975-2017 (SIAB7517),[4] a 2% random sample of the Integrated Employment Biographies (IEB). The IEB contain the administrative records of all employees liable to social security contributions in Germany and report precise to-the-day information on individual employment status and the daily wage received.[5] Starting in the year 1975 for West Germany and in 1992 for East Germany, the SIAB records the entire employment histories of approximately 1.8 million individuals. Since 1999, the SIAB additionally reports episodes in so-called 'minijobs' (also referred to as 'marginal employment'), which are not subject to social security contributions and pay a low wage not exceeding a certain threshold ('minijob' threshold). For our analysis, we determine each individual's main employment status on June 30 of a given year. This status is then assigned to one of three employment categories: 'full-time', 'part-time' and 'non-participation'. The categories of 'full-time' and 'part-time' refer to employment subject to social security contributions, while the category of 'non-participation' refers to non-

[4] We use the Scientific Use File supplied by the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB), see Antoni et al. (2019) for the data documentation.

[5] Wages are only reported up to the social security contributions ceiling, resulting in right-censoring that affects roughly 5 to 12% of all wage observations in each year. We impute these wages following Gartner (2005), as is common practice in the literature.

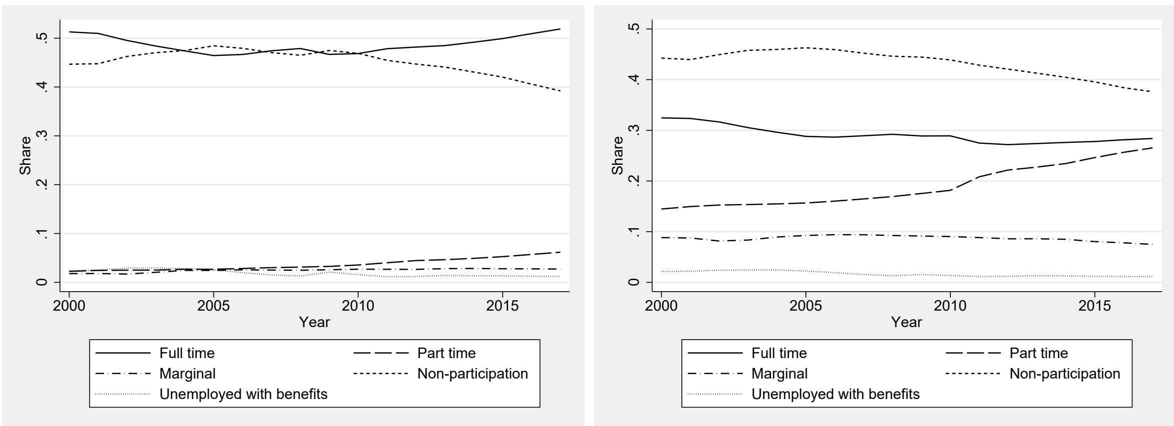
participation in any form of employment subject to social security contributions.

The SIAB is widely considered the most informative data source on employment and wages in Germany. Its administrative nature and large sample size make it an obvious choice for the given application, particularly because it is less prone to limitations typically associated with survey data. Since the reported wages serve as the basis for the calculation of pensions within the German social security system, the wage and employment data are necessarily free from measurement error. Similarly, attrition is not a concern since an individual's entire employment history is tracked as soon as the person's first employment episode is recorded. It thus follows that in the event of a gap occurring before, after, or between two adjacent employment episodes, the individual in question cannot have been employed subject to social security or in a minijob during this time frame. In our selection model, such gaps are generally labelled as 'non-participation'. It is important to mention that the data generally do not allow us to distinguish between different forms of non-participation in regular employment. Individuals in the pool of 'non-participants' may be unemployed or not active in the labour market, including those who are currently in education or on parental leave. In fact, they may work for pay, but are not liable to social security contributions. This applies to civil servants, the self-employed and those in undeclared work or work abroad. Figure B.1 in appendix B shows that between 4 and 5% of the German workforce are civil servants. Moreover, self-employment is much more prevalent among men (about 12.5 to 14% over the observation period vs. 7.5% for women).

For our main selection models we use a pooled sample of all years from 2000 to 2017, but we estimate the LGR separately for men and women. We construct our sample in the following way. An individual is included in the sample as soon as we observe a single employment spell (full-time, part-time, or marginal part-time) in any of the years 1975 to 2017. Furthermore, individuals are only included for periods in which they are aged between 20 and 60 years. For periods during which the individual is not observed in regular full-time or part-time employment,[6] we assign them the 'non-participation' status. Marginal employment is also assigned to the 'non-participation' state as we only distinguish between regular full-time, regular part-time, and non-participation in regular

[6] Throughout this paper, we refer to employment subject to social security contributions as 'regular' employment.

employment. Note that our sample excludes individuals who never participated in one of the employment forms registered in our data. This includes individuals who never actively participated in the labour market or who always worked as civil servants or self-employed. The broad definition of sample membership in our analysis, which considers an individual to be a sample member as soon as there is one observed employment spell at any time, is motivated by the fact that it is only in this way that we can capture many forms of female non-participation, which often takes the form of employment interruptions due to family-related reasons. However, it should be noted that our data does not include direct information on events such as childbirth, transitions to self-employment/civil service, or general household information.



(a) Men **(b)** Women

Figure 3.1 – Evolution of employment shares
Note: Full-time and part-time refer to employment subject to social security contributions.
(Source: SIAB v7517, own calculations)

Figure 3.1 provides an overview of employment trends in our final sample. The period we study spans from 2000 to 2017, including a time characterised by rising unemployment peaking in 2005, followed by a recovery phase that remained uninterrupted by the global financial crisis. This ultimately led to an unprecedented employment boom after 2010. The diversity of business cycle phases within our time frame makes it especially suitable for analysing the impact of shifting labour market conditions on employment sorting. The figure displays the shares of persons in full-time, part-time, marginal employment and non-participation as well as of those currently unemployed with regular unemployment benefits ('Arbeitslosengeld I').[7] The left panel shows that male full-time

[7]The figures for unemployment and marginal employment are only presented for descriptive purposes. In our analysis, we subsume these labour states in the 'non-participation' category.

employment shares were at their lowest point in 2005 – where both male unemployment and non-employment were the highest –, whereas they steadily increased after 2008. Likewise, the growth in male part-time employment picked up in 2010, while non-employment declined significantly.

The evolution of female employment shares followed a somewhat different pattern. The right panel shows that, for women, the employment boom was particularly driven by an increase in part-time employment, with its share rising sharply after 2010, while full-time employment for women saw only a modest recovery over the same period. By 2017, the shares of regular part-time and full-time employment had nearly converged. Marginal part-time employment was much more common among women, though its share remained stable at just under 10% throughout the observation period. Similar to men, non-employment among women declined significantly between 2005 and 2017. These shifts in employment shares for both men and women likely affected the composition of the employed and non-employed populations, potentially influencing inequality measures like the gender pay gap. Our analysis seeks to identify and adjust for such compositional changes in both genders.

Recall that the LGR selection model comprises three equations: the outcome equation (by distribution threshold), the selection equation, and the sorting equation (also by threshold). The dependent variable in the outcome equation is a binary for receiving a log real daily wage below a particular threshold $y \in \mathcal{Y}$ in a given employment state. An indicator of this employment state, in turn, serves as the dependent variable in the selection equation. Since we carry out separate estimations for full-time and part-time employment, the pools of selected and non-selected individuals differ depending on the employment type considered. The two groups need to be chosen carefully since their exact definitions bear important implications for the estimated latent distributions. In our model of full-time employment, we include those in part-time, marginal part-time and non-participation into the non-selected group, such that the resulting wage offer distribution is one that would result if all individuals in the population received full-time wages. In contrast, we exclude full-time workers from our analysis of part-time wages, considering only individuals in marginal part-time employment and those not participating in the regular labour force as the non-selected group. Consequently, the latent distribution of part-time wages reflects what would occur if all workers in the residual group of non-full-time employees were to receive part-time wages. This specification was chosen to avoid creating an inconsistent comparison group for part-time employees that would arise if both full-time workers and non-employed individuals were classified as non-selected in part-time employment.

As covariates determining the wage offer, we include six categories of age (25-29, 30-34, ..., 55-60 vs. 20-24 years), second-degree polynomials of work experience in full-time, part-time and marginal employment measured in years as well as indicators for zero experience in each of the three categories, regional (state-level) dummies, an indicator for German nationality, and the following categories of educational attainment: (1) lower/middle secondary education only (omitted), (2) lower/middle secondary education and completed vocational training, (3) upper secondary education ('Abitur') only, (4) upper secondary education and completed vocational training, (5) degree from a university of applied sciences ('Fachhochschule') and (6) university degree. Tables B.1 (full-time models) and B.2 (part-time models) in appendix B.3 present summary statistics for our selected and non-selected groups by gender. To allow for time effects in a flexible way, the outcome equation also includes indicators for the years 2001 to 2017. Similarly, a full set of year dummies is included in the sorting equation, which thereby determine the sign and magnitude of unobserved selectivity across years and distributional thresholds. A complete overview of the covariates used in each of the three equations is given in table B.3 in appendix B.

As discussed in section 3.3, and as in the vast majority of contributions in the literature, we leverage instrumental variables in order to estimate selectivity.[8] Suitable instruments for selection can be hard to find depending on the nature of the data used. Instrumental variables that have been used in the literature include the number of small children in the household, often in combination with marital status (Albrecht et al., 2009; Maasoumi and Wang, 2019; Mulligan and Rubinstein, 2008), husband's income (Buchinsky, 2001), and potential out-of-work income (Arellano and Bonhomme, 2017; Elass, 2024). The first

[8]A notable and very recent exception is the approach suggested by Chen et al. (2024), who adapt the Arellano and Bonhomme (2017) method by replacing the binary selection equation with a censored equation of hours worked, rendering the use of instruments redundant. Since our administrative data only includes full- and part-time status, but not hours worked, this approach is not feasible in our application. D'Haultfœuille et al. (2018) also do not invoke exclusion restrictions, but base identification on extreme wage observations. Such observations are not available in our data due to censoring. Both approaches also do not allow one to study the strength of selection at different points in the distribution.

two options are not feasible in our analysis because our data lack household information. Additionally, we cannot utilise potential out-of-work income, as it is closely tied to earnings potential, i.e., prior earnings, and household characteristics within the German unemployment insurance system. Moreover, family-related instruments suffer from the essential drawback that they typically apply better to women than to men, while one of our objectives is to describe selectivity patterns for both men and women.

Instead, we opt for a group-instrument strategy that leverages variation in employment dynamics across groups of workers and over time. To the best of our knowledge Fitzenberger and de Lazzer (2022) was the first study using this type of instrumentation strategy, which addresses the fact that administrative data typically lack other information that can be used for instrumentation. In our application, we exploit employment dynamics within worker cells defined by sex, year, age, educational attainment, and local labour markets ("Raumordnungsregionen", ROR) in order to explain varying selection into full-time and part-time employment at the level of the individual. We use 15 different indicators which we compute by worker cell: full-time share, part-time share, minijob share, annual first differences of these shares, as well as the full matrix of annual transition rates between full-time employment, part-time employment and non-participation. The reader is referred to table B.3 in appendix B for more details. As additional instruments, we include indicators that mark persons of three different age groups who are likely to be students in the given year (i.e., persons who did not hold a college degree in that year, but are reported to have one at a later stage in their employment history).

Our instrumental variables are valid provided that the employment dynamics represented by the above indicators influence wages only with a sufficient time lag. We assume this to be the case in the German labour market, where wage rigidities prevent short-term adjustments of wages (e.g., Bauer et al., 2007). Wage contracts typically span multiple years, and the remuneration of new employees is usually aligned with the rates paid to existing staff. Additionally, collective bargaining occurs at the level of the industry or region, with a notable lag, and is unlikely to respond directly to developments within the disaggregated worker cells that we have considered. Apart from variation between worker groups and regions, our instruments represent rich variation over time reflecting the different phases of the business cycle (rising unemployment 2000-2005, stagnation

2006-2009, boom 2010-2017; see figures B.2 and B.3 in appendix B). We consider the exclusion restriction underlying our instruments not to be stronger than the assumptions made in other studies in the literature. For example, the presence of children and a husband's income have been criticised by Huber and Mellace (2014), while Blundell et al. (2007) question the validity of out-of-work income.

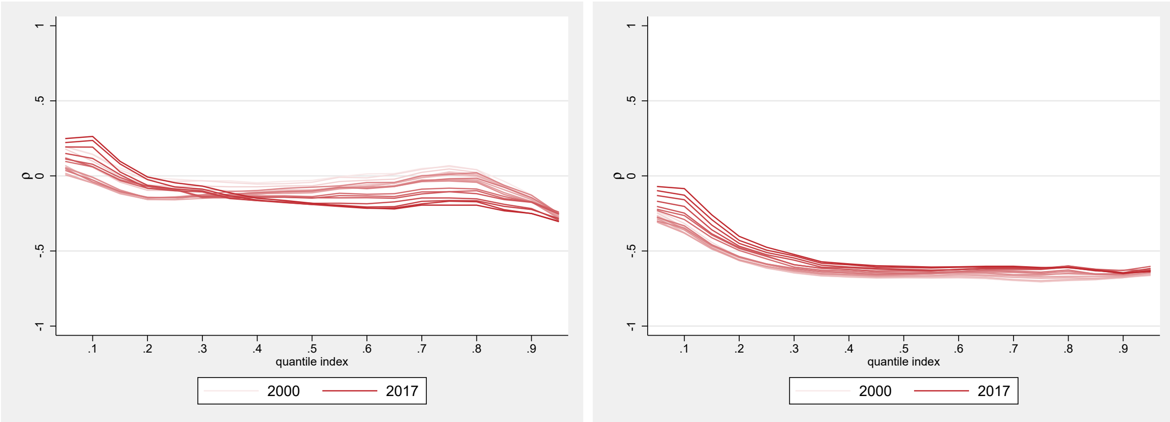
3.5 Empirical results

We present three sets of results, both for full-time and for part-time wages. First, we describe male and female patterns of unobserved selectivity and how these evolved over time. Second, we decompose the gender gap in latent wages into their main sources of observable worker characteristics and returns to these characteristics. Finally, we decompose differences in *observed* wage distributions between men and women and across time periods into the different components implied by our rich distributional selection model.

3.5.1 Selection into full-time employment

Figure 3.2 presents the estimated sorting parameters for full-time employment against all other employment states, where the sorting equations $\rho(x'\hat{\delta}(y))$ include a constant and dummies for the years 2001 through 2017. As in Chernozhukov et al. (2025), we show the sorting parameters at wage thresholds defined by a fine grid of percentiles of the distribution of log real wages with indexes {0.05, 0.10, ..., 0.90, 0.95} computed in the pooled sample of men and women. The change in colour saturation, moving from pale to more vibrant red hues, reflects the progression of time from the year 2000 to 2017.

As a first observation, both graphs reveal sizeable heterogeneities in sorting across the wage distribution for both genders, where the estimated selection bias increases in the quantile index by absolute value towards the top for women and changes its sign for men. This challenges the results obtained from other methods that only estimate one single sorting parameter for the whole distribution. Second, figure 3.2 shows non-negligible female *and* male selection, which has thus far received limited attention in the literature. Third, the observed selection patterns become somewhat more heterogeneous over time,



(a) Men **(b)** Women

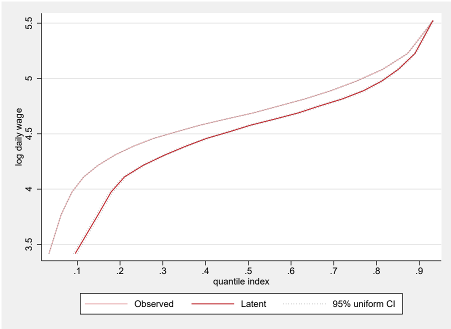
Figure 3.2 – Estimated parameter of sorting into full-time work
Note: Increasing dark red for more recent years.
(Source: SIAB v7517, own calculations)

most notably for men: while the selection bias tended to be relatively small across the distribution directly after the millennium, sorting became larger in magnitude as well as increasingly heterogeneous towards more recent years. Similarly, there is a tendency for increasingly heterogeneous sorting from 2000 to 2017 for women – as implied by the steeper $\rho(x'\hat{\delta}(y))$ curve in the right panel – but the general pattern does not change as much as it does for men.

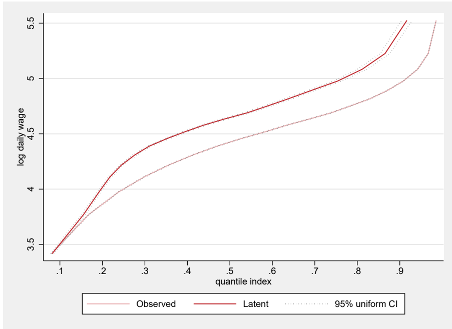
Overall, figure 3.2a suggests that in the early 2000s (a period characterised by high unemployment), men were positively selected at the bottom as well as between the 60th and 80th percentiles and almost non-selected in other parts of the wage distribution. In contrast, male selection is only positive at the bottom, but negative in other parts towards the more recent years that have witnessed record employment levels. This intertemporal pattern supports the hypothesis that the composition of the workforce in terms of unobservables becomes less favourable during periods of high employment as individuals with less favourable unobservables enter the labour market (Dolado et al., 2020) – or individuals with less favourable unobservables are less likely to exit the labour market (Riphahn and Schrader, 2020). Empirically, a similar result was obtained by Fitzenberger and de Lazzer (2022) who conclude that the unemployed are negatively selected in times of high employment. A potential explanation for male negative selection in our sample is the fact that our non-selected group partly contains the self-employed and civil servants (see section 3.4), who may be positively selected against the employed subject to social security. In figure 3.2a, we observe *positive* unobserved selectivity for men at the very bottom of the distribution. A potential explanation for positive selectivity in this part of the distribution is the generous German social safety net, which may disincentivise employment for individuals with the lowest wage offers.

In contrast to men, women experience negative selection across the entire distribution in all years. Selection is close to zero at the bottom, especially in the later years (figure 3.2b). Negative female selection has been found both by Elass (2024) for Finland and the United Kingdom over a period covering the Great Recession, and by Maasoumi and Wang (2019) for the United States until the 1990s. Likewise, our results align with those of Chernozhukov et al. (2025) who document increasingly negative sorting of British women towards the top of the distribution. From a theoretical point of view, Ermisch and Wright (1994) argue that negative sorting can be plausible when there is a high positive correlation between wage offers and reservation wages. One possible explanation for this is assortative matching on the marriage market, with couples becoming increasingly homogeneous in terms of education, income, and likely also unobserved ability (see Calvo et al., 2024, for evidence on assortative matching in Germany). After the birth of a child, in particular, assortative matching may result in high-potential women taking on the bulk of care work because their high-potential partner's income is sufficient to maintain household income, while women with less successful partners face the necessity to contribute to family income. Negative sorting of married women towards the top of the distribution is also consistent with Germany's tax and transfer system, which overproportionally favours single-income marriages and marriages with a high intercouple income gap since the secondary earner often faces high marginal tax rates (Bach et al., 2013). According to Figure 3.2b, female selectivity became somewhat less negative in more recent years. This may be due to improvements in public childcare and changing social norms which facilitated the return to full-time work after childbirth (Geyer et al., 2015).

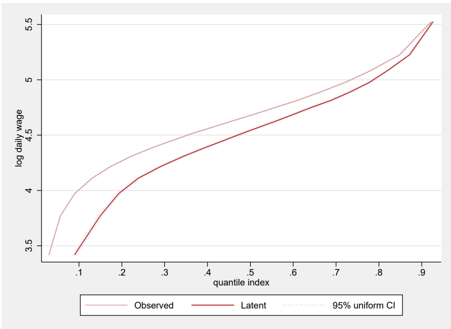
As explained in section 3.3, our estimates of selection patterns can be used to recover the latent wage offer distributions for men and women (eq. (3.7)), which can be compared to the observed distribution (eq. (3.10)). Figure 3.3 presents these comparisons separately for the period 2000-2005 and 2012-2017 in order to assess potential changes over time. For men, the left panels show that the observed distribution of full-time wages first-order stochastically dominates the latent distribution (i.e., observed wages are higher than



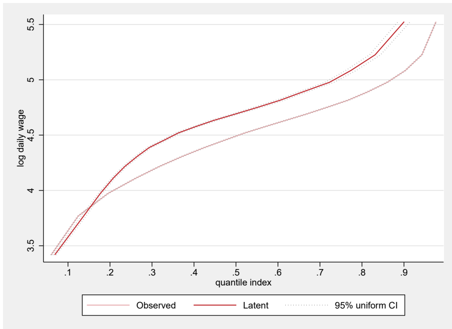
(a) Male latent vs. observed full-time wages, 2000-2005



(b) Female latent vs. observed full-time wages, 2000-2005



(c) Male latent vs. observed full-time wages, 2012-2017



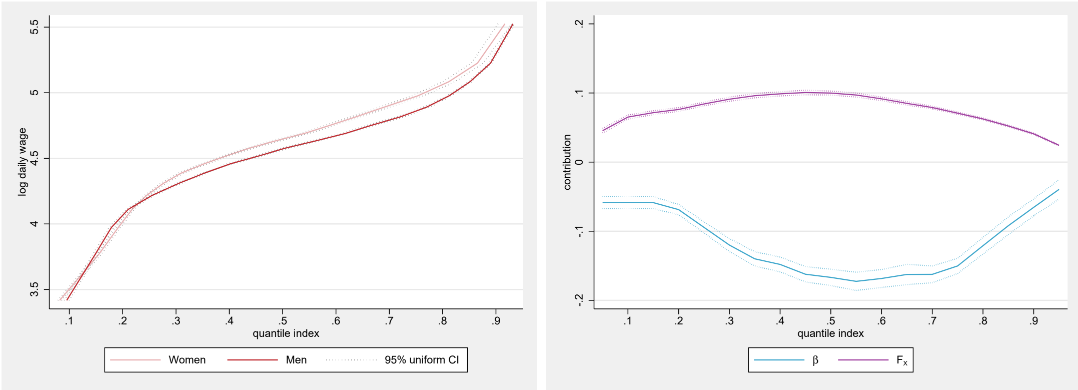
(d) Female latent vs. observed full-time wages, 2012-2017

Figure 3.3 – Latent vs. observed distributions of full-time wages, 95% uniform confidence bands
(Source: SIAB v7517, own calculations)

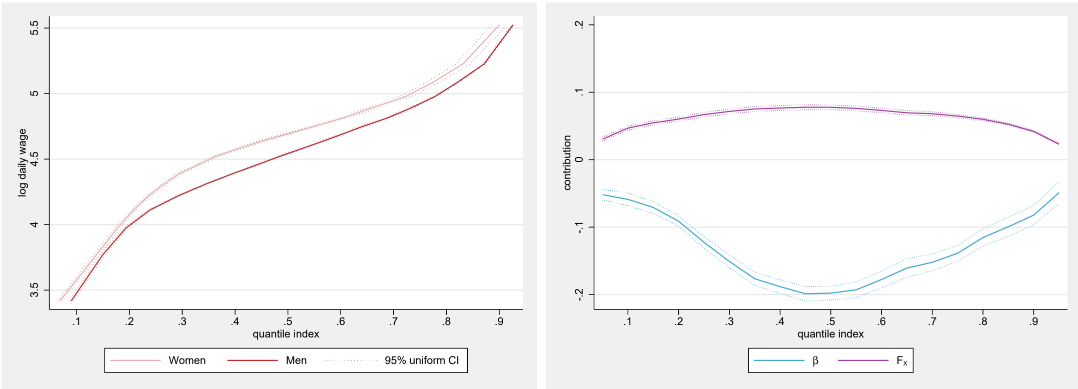
latent wages). This may be surprising given the negative selection pattern for men above the 30th percentile (figure 3.2a). Recall however, that the latent distribution of full-time wages (eq. (3.7)) also incorporates observables for men who do not select themselves into full-time work. These are much less favourable than for men who work full-time (table B.1). In particular, men participating in regular full-time employment are more favourably selected with respect to work experience, education and nationality.

For women, the right panels of figure 3.3 paint a different picture. With the exception of the lowest income levels, where the latent and observed distributions coincide, the latent full-time wage is higher than the observed wage. This is in line with their distinctively negative selection on unobservables (figure 3.2b), which dominates the positive selection on observables (table B.1). For both men and women, latent full-time wage distributions are relatively stable, despite some temporal change in selection on unobservables (figure 3.2).

Next, we decompose the gender gap in *latent* full-time wages between men (group 0)



(a) Female vs. male latent wages, 2000-2005 **(b)** Individual contributions, 2000-2005



(c) Female vs. male latent wages, 2012-2017 **(d)** Individual contributions, 2012-2017

Figure 3.4 – Decomposition of differences in latent full-time, wage distributions between men and women, 95% uniform confidence bands
(Source: SIAB v7517, own calculations)

and women (group 1) into a contribution due to differences in wage structures and a contribution due to differences in the composition of worker characteristics, i.e.,

$$F_{Y^*[1,1]} - F_{Y^*[0,0]} = \underbrace{\left[F_{Y^*[1,1]} - F_{Y^*[0,1]}\right]}_{\text{wage structure}} + \underbrace{\left[F_{Y^*[0,1]} - F_{Y^*[0,0]}\right]}_{\text{composition}}, \tag{3.13}$$

where

$$F_{Y^*\langle j,k \rangle}(y) = \int \Phi(-x'\beta_j(y))dF_{X_k}(x) \tag{3.14}$$

with $\beta_j(y)$ denoting the coefficients in the wage equation for group $j \in \{0,1\}$, and F_{X_k} the distribution of characteristics of group $k \in \{0,1\}$.[9]

Perhaps surprisingly, the latent full-time wage distribution of women is slightly above that of men (left panel in figure 3.4). The contributions to the difference between the

[9]Our conclusions are unchanged when the order of the decomposition is reversed.

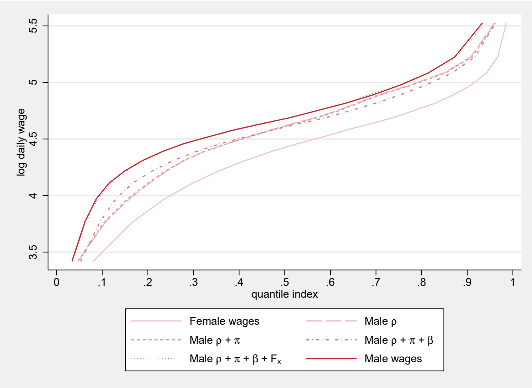
latent CDFs of men and women (as defined in (3.13)) are shown in the right panels of figure 3.4. In order to facilitate interpretation, we switch the signs of the contributions in the graphs, so that a positive contribution means that the wages of group 1 (i.e., women) are lifted upwards (equivalently, their CDF is pulled downwards) if a given decomposition factor is replaced by its group 0 (i.e., male) counterpart. The right-hand panels of figure 3.4 show that there are two countervailing effects: applying male wage returns to women pulls down their wage distribution, while applying the more favourable distribution of male observables (table B.1) shifts them upwards. The observation that returns to characteristics are less positive for men than for women may be surprising. However, recall that these are returns *corrected for selectivity*. As described above, selectivity on unobservables is distinctively negative for women (figure 3.2b). This downward biases estimated wage returns when sample selection is not corrected for. An interpretation of this finding is that, if institutional regulations limit observable pay differences between two worker groups (as represented by returns uncorrected for selectivity), this may mask higher actual returns for the more negatively selected worker group.

We now turn to differences in *observed* full-time wages between men and women. Note that this analysis differs from the previous one in that it only refers to men and women *observed in full-time employment* (while the previous analysis referred to the full population of *all* men and women). Similar to Chernozhukov et al. (2025), we decompose these differences into five components implied by the distributional selection model: (i) sorting on unobservables, (ii) differences in employment structure, (iii) differences in the wage structure, (iv) differences in observables, and (v), differences in labour market dynamics.[10]

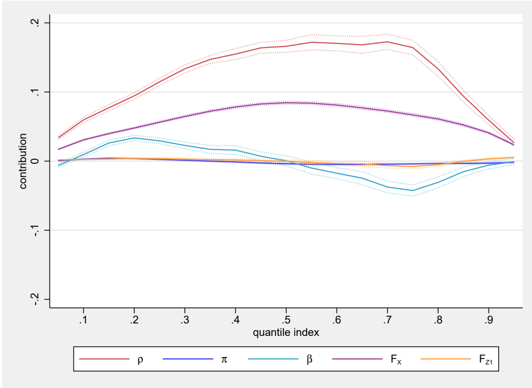
The decomposition is given by

$$\begin{aligned}
F_{Y[1,1,1,1,1]} - F_{Y[0,0,0,0,0]} &= \underbrace{\left[F_{Y[1,1,1,1,1]} - F_{Y[0,1,1,1,1]}\right]}_{\text{selection sorting } (\rho(X'\delta(y)))} + \underbrace{\left[F_{Y[0,1,1,1,1]} - F_{Y[0,0,1,1,1]}\right]}_{\text{selection structure } (\pi)} \\
&+ \underbrace{\left[F_{Y[0,0,1,1,1]} - F_{Y[0,0,0,1,1]}\right]}_{\text{wage structure } (\beta(y))} + \underbrace{\left[F_{Y[0,0,0,1,1]} - F_{Y[0,0,0,0,1]}\right]}_{\text{composition } (F_X)} \\
&+ \underbrace{\left[F_{Y[0,0,0,0,1]} - F_{Y[0,0,0,0,0]}\right]}_{\text{labour market dynamics } (F_{Z_1})}, \qquad (3.15)
\end{aligned}$$

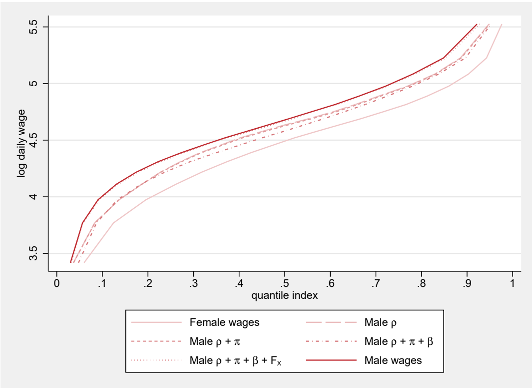
[10] Compared to Chernozhukov et al. (2025), we add component (v) as we can readily swap the values for the labour market instruments between the genders.



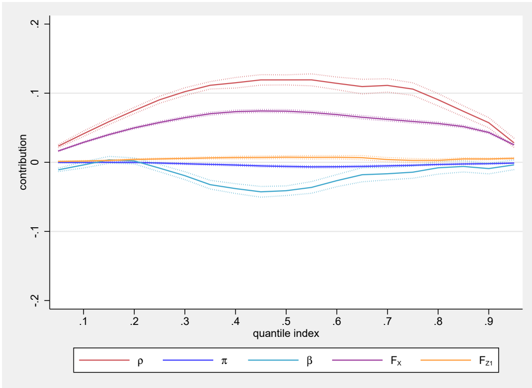
(a) Observed and counterfactual distributions, 2000-2005



(b) Individual contributions, 2000-2005



(c) Observed and counterfactual distributions, 2012-2017



(d) Individual contributions, 2012-2017

Figure 3.5 – Decomposition of differences in observed full-time wage distributions between men and women, 95% uniform confidence bands
(Source: SIAB v7517, own calculations)

where

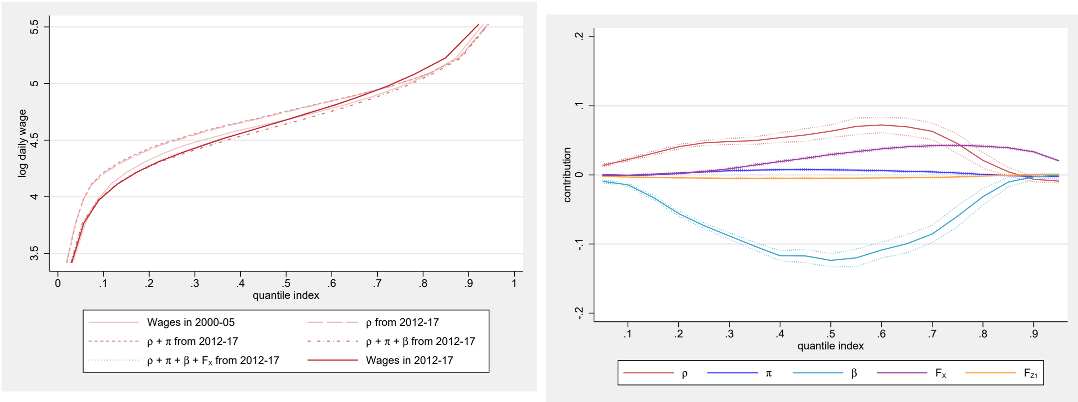
$$F_{Y\langle t,s,r,k,l\rangle}(y) = \frac{\int \Phi_2(-x'\beta_s(y), z'\pi_s; -\rho(x'\delta_t(y)))dF_{Z_{1l},X_k}(z_1, x)}{\int \Phi(z'\pi_s)dF_{Z_{1l},X_k}(z_1, x)}. \tag{3.16}$$

A common finding is that observed male wages are higher than those for females. This is also confirmed in our case (bold and light red lines in figure 3.5). Decomposition (3.15) starts with the female full-time wage distribution $F_{Y[1,1,1,1,1]}$ and assigns the much less negative sorting behaviour of men to women. This shifts the wage distribution upwards, as shown in figure 3.5a. Again, we represent this as a positive contribution in the right-hand figure 3.5b. In the next decomposition step, we additionally apply the male selection structure effects π. This does not change the distribution in a meaningful way. Next, we assign the male wage structure, which raises female wages in the lower half but dampens them in the upper half of the distribution (blue line in figure 3.5b). In figure 3.5d for the period 2012 to 2017, this contribution is mostly negative, i.e., favourable for females. Also adding the male observables distribution leads to a significant upward shift of wages, which is not surprising given that male full-time employees have better observables than their female counterparts (table B.1). Finally, switching female labour market dynamics (as represented by the instruments) to those of men does not lead to any noticeable changes in figure 3.5a for the period 2000 to 2005. However, there is a slightly positive contribution for 2012 to 2017 owing to the fact that the development of full-time employment was more favourable for men than for women over this period (figure 3.1).

Taken together, our decomposition suggests that a significant portion of the observed gap in full-time wages between men and women can be attributed to the less negative unobserved selectivity of men compared to women. Additionally, a substantial part of the gap arises from men in full-time employment having better observable characteristics than women. When comparing the periods from 2000-2005 and 2012-2017, the wage gap between men and women has substantially narrowed. As shown in the right-hand panels of figure 3.5, this reduction is largely driven by declining differences in unobserved selectivity – women's selectivity became less negative, while men's selectivity turned more negative (figure 3.2). Another, smaller contribution comes from returns to observable characteristics, which became more favourable for women than for men in 2012-2017 compared to 2000-2005 (blue lines in figures 3.5b and 3.5d).

Finally, we use decomposition formula (3.15) to examine changes in full-time wages over time. For this, we estimate our selection model separately for the periods 2000-2005 and 2012-2017 and by gender. Figure 3.6 displays the results for men, where the left panel successively moves from the factual distribution in 2000-2005 (group 1) to that of 2012-2017 (group 0). It shows that the lower half of the 2012-2017 distribution of real full-time wages lies below that of 2000-2005, but that the upper half lies slightly above it. This is in line with earlier contributions showing that the development after 2005 was characterised by a stagnation or decline of real wages in the lower half of the distribution (Baumgarten et al., 2020; Biewen and Sturm, 2022; Dustmann et al., 2014). Figure 3.6b suggests that this was the result of a complex mix of different effects: declining real wage returns in the middle of the distribution, improved selection in the lower and middle part of the distribution and better observables at the top of the distribution. The strong pattern of declining real returns is consistent with evidence in Dustmann et al. (2014)

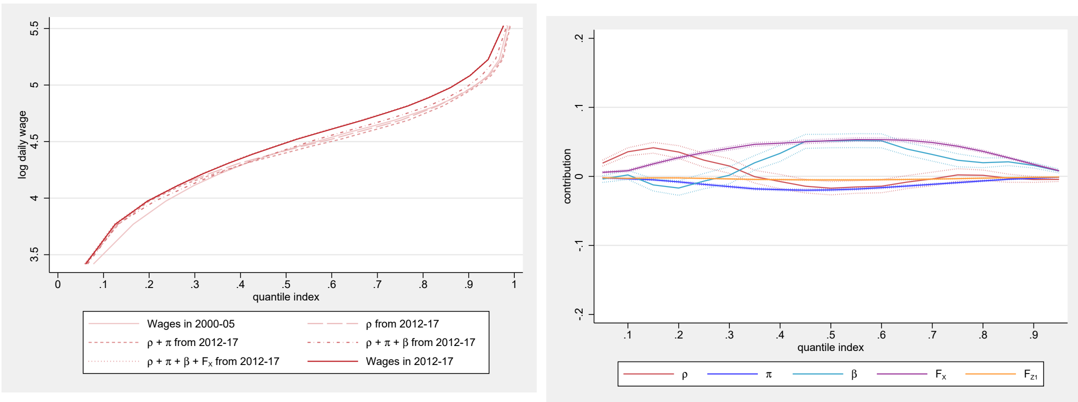
who concluded that wage setting in Germany became significantly more flexible after 2005.



(a) Male observed and counterfactual wage distributions

(b) Individual contributions

Figure 3.6 – Decomposition of differences in observed male full-time wage distributions between 2000-2005 and 2012-2017, 95% uniform confidence bands
(Source: SIAB v7517, own calculations)



(a) Female observed and counterfactual wage distributions

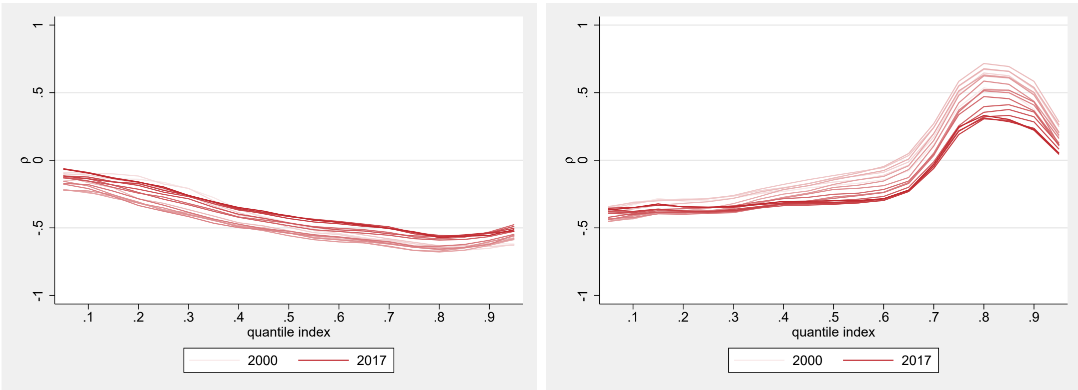
(b) Individual contributions

Figure 3.7 – Decomposition of differences in observed female full-time wage distributions between 2000-2005 and 2012-2017, 95% uniform confidence bands
(Source: SIAB v7517, own calculations)

Unlike men, women experienced moderate real wage gains across all parts of the distribution from 2000-2005 to 2012-2017 (figure 3.7a). Figure 3.7b shows that this is the result of more favourable observables, increased real wage returns in the middle of the distribution, and, to a smaller extent, less negative selection at the lower end of the distribution. The strongest effect comes from improved observables, in particular from a higher proportion of women with a university degree and higher levels of full-time work experience (table B.1).

3.5.2 Selection into part-time employment

We repeat the analysis for part-time wages. Recall from section 3.4 that we model selection into part-time employment for the residual population of men and women who do not work in regular full-time. For these individuals, we consider participation in part-time employment vs. all other labour market states.

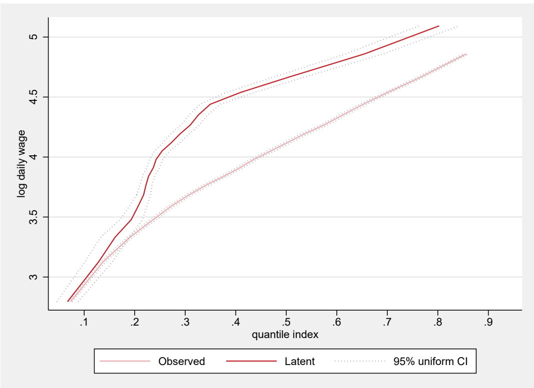


(a) Men **(b)** Women

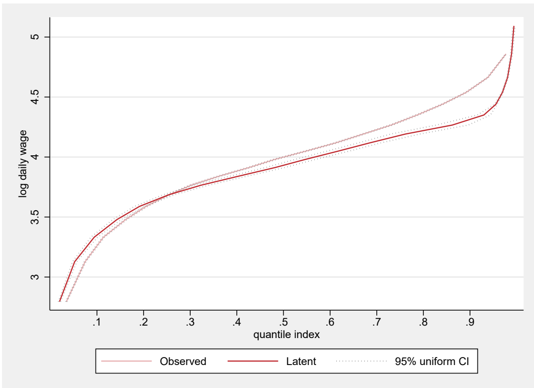
Figure 3.8 – Estimated parameter of sorting into part-time work
Note: Increasing dark red for more recent years.
(Source: SIAB v7517, own calculations)

For men, figure 3.8a suggests negative unobserved selectivity in part-time work, which became less negative towards more recent years. The phenomenon of negative selection in part-time employment can be explained by the fact that the residual population of men who do not work in regular full-time employment comprises a significant proportion of individuals who at some point moved into self-employment, civil service or work abroad. Unfortunately, we cannot differentiate these cases with our dataset. Compared to this group, men in part-time employment may be a negative selection with respect to unobservables, especially for higher levels of pay. Out of the residual group of men not working full-time, only a small fraction actually chose to take up regular part-time employment (table B.2). However, this fraction significantly increased over time. Both in 2000-2005 and 2012-2017, around 1.3 million men in our sample did not work full-time. Of these, around 150 thousand worked part-time in 2012-2017, compared to around 75 thousand in 2000-2005. The fact that part-time work became more common among men may have also contributed to the decline in negative selectivity as shown in figure 3.8a.

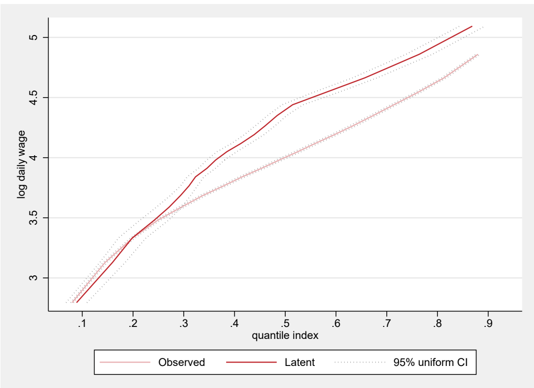
The corresponding figure 3.8b for women reveals a complex sorting pattern into part-time work with a change of sign between the 0.6 and 0.7 quantile indices. A potential explanation for the observed pattern may again be related to assortative matching. In the lower to middle part of the distribution, even women with less favourable unobservables would often be forced to contribute to household income because their partners have low income. On the other hand, women with more favourable unobservables may generally be disincentivised by the structure of the German tax system which favours large earnings differentials between partners. These disincentives are particularly strong in the upper part of the distribution, so that only women with the highest wage offers decide to take up additional part-time employment (Bick and Fuchs-Schündeln, 2017). This pattern of positive selection at the top weakened in later years, possibly because female part-time work became increasingly common (figure 3.1), or because changes in social norms and public childcare made employment after childcare more acceptable.



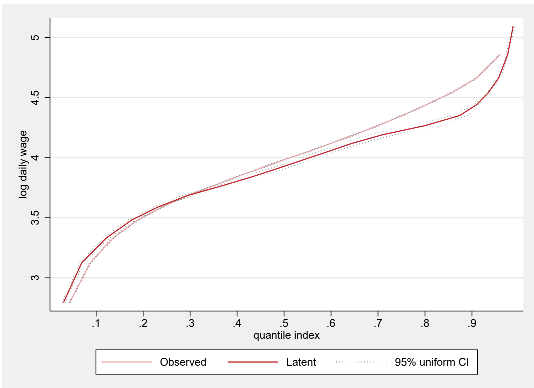
(a) Male latent vs. observed part-time wages, 2000-2005



(b) Female latent vs. observed part-time wages, 2000-2005



(c) Male latent vs. observed part-time wages, 2012-2017

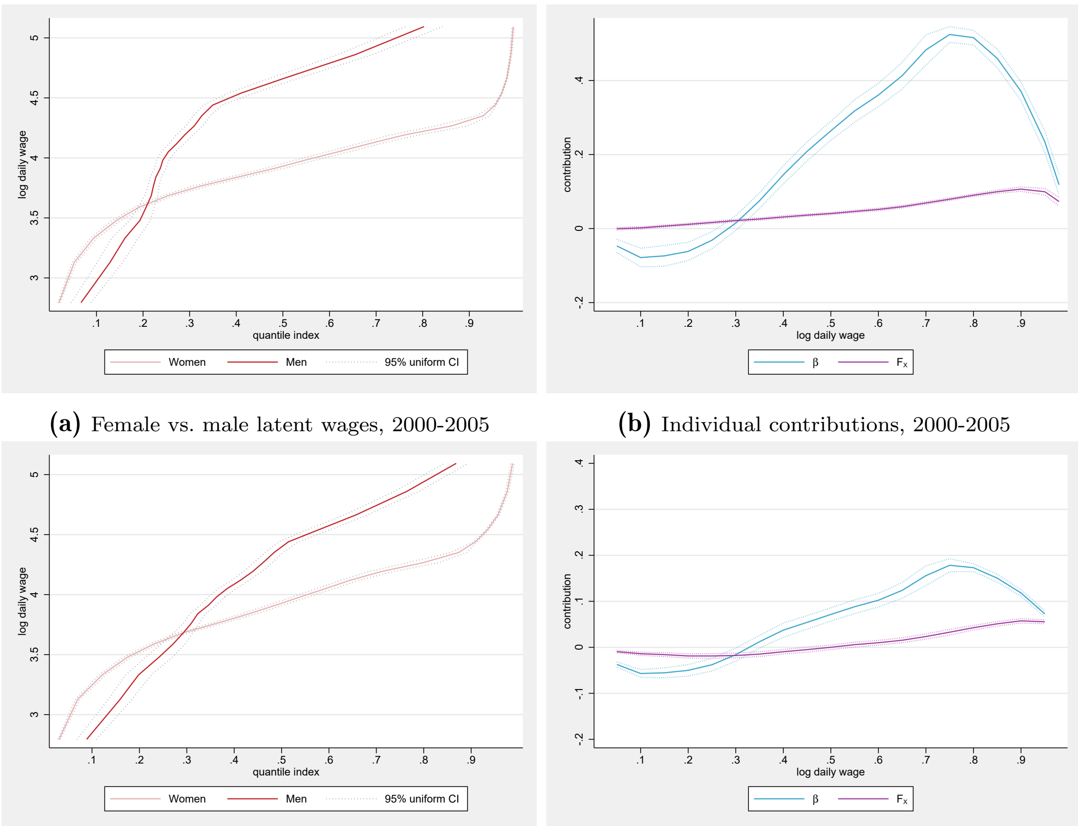


(d) Female latent vs. observed part-time wages, 2012-2017

Figure 3.9 – Latent vs. observed distributions of part-time wages, 95% uniform confidence bands

(Source: SIAB v7517, own calculations)

Figure 3.9 compares latent and observed part-time wage distributions for men and women. For men, the pronounced negative selection on unobservables (figure 3.8a) clearly dominates the somewhat positive selection on observables (table B.2) so that latent wages lie substantially above observed part-time wages (left panel of figure 3.9). For women, negative selection on unobservables balances with positive selection on observables up to around the 70th percentile, above which positive selectivity on unobservables pushes observed wages above latent wages (right panel of figure 3.9). For both men and women, differences between latent and observed part-time wage distributions declined in recent years, driven by differences in unobserved selectivity over time (figure 3.8).



(a) Female vs. male latent wages, 2000-2005

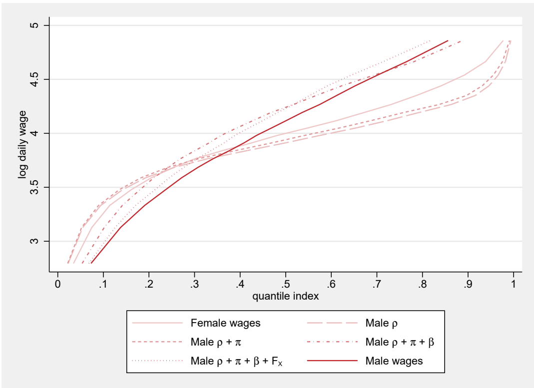
(b) Individual contributions, 2000-2005

(c) Female vs. male latent wages, 2000-2005

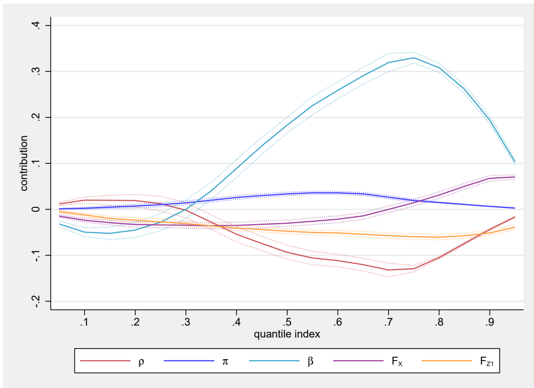
(d) Individual contributions, 2012-2017

Figure 3.10 – Decomposition of differences in latent part-time wage distributions between men and women, 95% uniform confidence bands
(Source: SIAB v7517, own calculations)

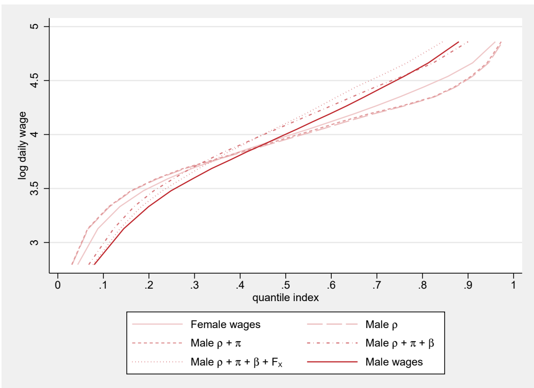
Figure 3.10 presents the gender gap in *latent* part-time wage distributions. Again, these distributions are constructed by assigning *all individuals* (out of the residual group of men and women not working in regular full-time) the part-time wages of the respective group. This leads to very large differences between men and women. As the right-hand side of figure 3.10 shows, the large differences are mainly due to the much better wage returns for men in part-time work compared to women, while differences in the distribution of observables play a more modest role. A potential explanation for this very pronounced wage structure effect is that our set of observables contains important characteristics such as age, education and experience but lacks more detailed characteristics of male and female part-time job profiles. It may be the case that the relatively small number of part-time men are employed in roles that require higher qualifications, as opposed to part-time women whose range of part-time employments tends to be much wider. The bottom panel of figure 3.10 again shows signs of convergence between male and female distributions, driven by a convergence in wage returns (blue lines in figures 3.10b and 3.10d). This convergence is a likely consequence of part-time work becoming more common among men, thereby expanding the range of job types typically available to men in part-time roles.



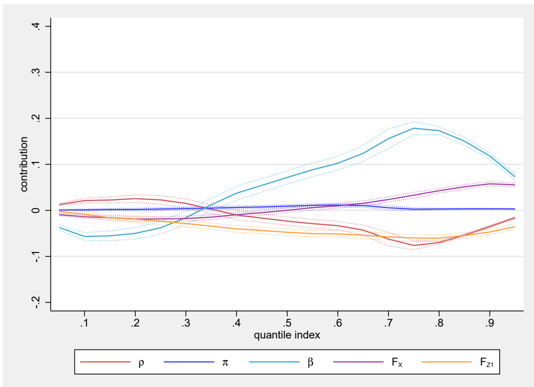
(a) Observed and counterfactual distributions, 2000-2005



(b) Individual contributions, 2000-2005



(c) Observed and counterfactual distributions, 2012-2017

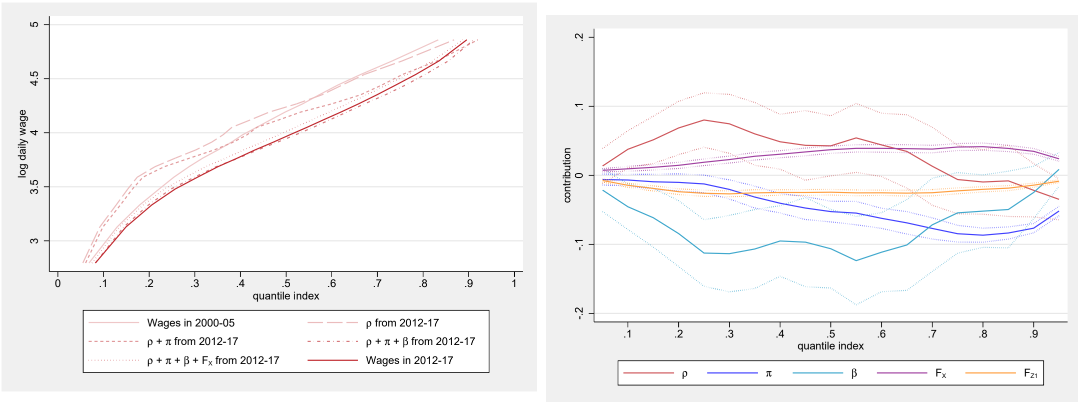


(d) Individual contributions, 2012-2017

Figure 3.11 – Decomposition of differences in observed part-time wage distributions between men and women, 95% uniform confidence bands
(Source: SIAB v7517, own calculations)

Figure 3.11 examines differences in *observed* part-time wage distributions between men and women (decomposition (3.15)). Male part-time wages dominate female part-time wages in the upper part of the distribution, while the reverse is true in the lower part of the distribution. Again, it turns out that part-time men in the upper half of the distribution enjoy much higher wage returns than part-time women (blue lines in figures 3.11b and 3.11d), which is the likely result of better job characteristics not controlled for in our set of observables. The effect of higher male wage returns is partly counteracted by more positive female selection on unobservables (red lines in figures 3.11b and 3.11d). Observed male part-time wages only dominate female part-time wages above the 40th percentile (left panel of figure 3.11). In the lower part of the distribution, women have better observables and face slightly better wage returns than men (blue and purple lines in figures 3.11b and 3.11d). Differences in labour market dynamics between men and women also play a small role, boosting observed part-time wages of women (orange line in figures 3.11b and 3.11d). The comparison between the upper and the lower panel of figure 3.11 again indicates convergence between male and female part-time wage distributions over time, mostly due to diminishing differences in wage returns (blue lines in right panel) and declining differences in selectivity on unobservables (red lines in right panel).

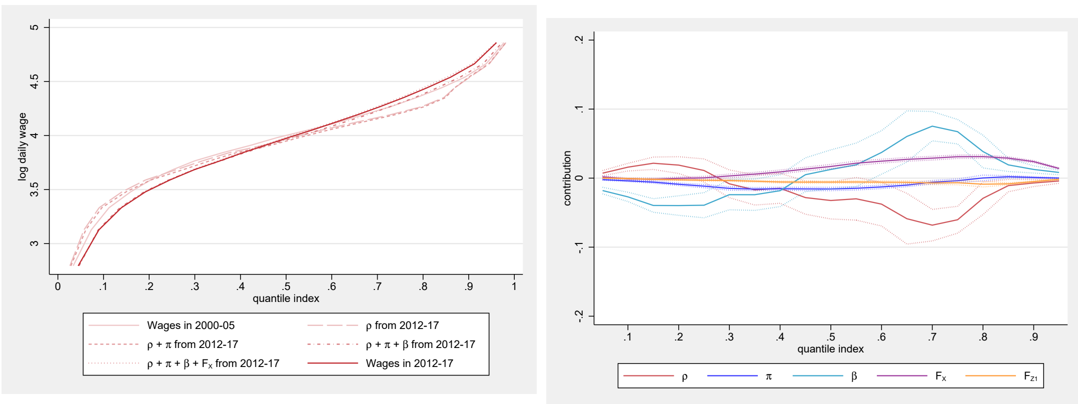
Our final analysis concerns the decomposition of the evolution in observed male and female part-time wage distributions over time (figures 3.12 and 3.13). Figure 3.12 for men indicates a deterioration of part-time wages for men between 2000-2005 and 2012-2017 (the distribution of 2000-2005 lies above that of 2012-2017, see fine and bold red lines in figure 3.12). The right panel of figure 3.12 suggests that, although there were improvements in selection on unobservables and selection on observables (red and purple lines in figure 3.12b), these were more than offset by a decline in part-time wage returns (blue line) and a selection structure effect (dark blue line in figure 3.12b). This deterioration in part-time wage returns is in line with the above conjecture that the proportionally massive expansion of male part-time employment between 2000-2005 and 2012-2017 involved a change in job profiles not captured by our observables. Furthermore, the selection structure effect, as introduced in Chernozhukov et al. (2025) (depicted in dark blue), indicates that male selection into part-time employment underwent a transformation whereby observable characteristics, specifically age, education, and experience, were increasingly matched with less favourable part-time roles, particularly at the upper end of the part-time distribution. The much lower statistical precision of the point estimates in figure 3.12b reflects the relatively small sample size of this group of workers (table B.2).



(a) Male observed and counterfactual wage distributions

(b) Individual contributions

Figure 3.12 – Decomposition of differences in observed male part-time wage distributions between 2000-2005 and 2012-2017, 95% uniform confidence bands
(Source: SIAB v7517, own calculations)



(a) Female observed and counterfactual wage distributions

(b) Individual contributions

Figure 3.13 – Decomposition of differences in observed female part-time wage distributions between 2000-2005 and 2012-2017, 95% uniform confidence bands
(Source: SIAB v7517, own calculations)

The analysis of female part-time wages over time is given in figure 3.13. Apart from some gains above the 60th percentile and some losses below, the female part-time distribution remained relatively stable between 2000-2005 and 2012-2017. Figure 3.13b suggests that the gains were the result of improved wage returns to observed characteristics (blue line), which overcompensated the decline in positive selection on unobservables (red line). Better observable characteristics, in particular, more work experience and a much higher proportion of workers with university degrees (table B.2), also contributed to higher female part-time wages in 2012-2017 compared to 2000-2005 (purple line in figure

3.13). The corresponding pattern is strikingly similar to that observed for males (figure 3.12), which suggests a general upgrading of part-time jobs in Germany. At the lower end of the female part-time wage distribution, there was an erosion of wage returns (blue line), which was not compensated by slight improvements in unobserved selectivity (red line in figure 3.13b).

3.6 Conclusion

Based on high-quality administrative data, this paper examines male and female wage distributions in Germany while flexibly accounting for selection on unobservables. Our findings suggest that male selectivity in full-time work is positive at the lower end of the distribution but becomes negative throughout the rest of it, a trend that intensified over time. This aligns with the hypothesis that the post-2012 full-time employment boom drew individuals with less favourable unobservables into the workforce. Positive selectivity at the bottom of the distribution may be explained by the generosity of Germany's social safety net, which discourages employment among those with very low wage offers. For women, full-time employment generally exhibits negative selectivity, potentially driven by assortative matching and household dynamics, where women with better unobservables may not need to contribute to household income. Recent improvements in public childcare and evolving social norms appear to have mitigated this negative selectivity by encouraging women's return to full-time work after childbirth.

A significant portion of the full-time gender wage gap is explained by differences in unobserved selectivity between men and women. Better observables of full-time men (especially work experience and education) explain another large share of the gap, while selectivity-corrected wage returns work in favour of women. Between 2000-2005 and 2012-2017, male full-time wages rose in the upper half of the distribution but declined in the lower half, reflecting wage erosion across broad segments of the distribution. We attribute this to wage restraints and increased wage flexibility (Dustmann et al., 2014). At the top of the distribution, improved selection on unobservables and better observables contributed to higher male wages in the later period. By contrast, female full-time wages grew between 2000-2005 and 2012-2017 across the whole distribution as a result of better observables, increased wage returns in the middle and less negative selection at the lower end of the distribution. Overall, the full-time wage gap between men and women significantly narrowed between 2000-2005 and 2012-2017, largely due to declining differences in unobserved selectivity and improved observables for women.

Our paper is among the first to provide an in-depth examination of unobserved selectivity in part-time employment. We find that men working part-time represent a negatively selected subset of the group of men not pursuing regular full-time employment. This may be explained by the fact that the group of men in our data not observed in full-time employment also include the self-employed and other individuals working for pay, who are not subject to social security contributions. In the past, only very few men worked part-time in Germany, making this group of workers a highly specific one. Male selectivity in part-time work became less negative as part-time employment among men has grown, although it remains much less common than for women. At the lower end of the wage distribution, negative selectivity is less pronounced, possibly due to the same social safety net effects observed in full-time employment. For women, part-time work exhibits a complex selectivity pattern turning from negative at the bottom to positive at the top. Our explanation for this pattern is that assortative matching forces women in the lower part of the distribution to contribute to household income, while women in the upper part only work if they receive very high wage offers. Recent declines in female part-time selectivity may reflect improved public childcare and shifting social norms supporting post-childbirth employment.

We find that male part-time wages are higher than those of females in the upper half but lower in the bottom half. This is due to higher wage returns in male part-time jobs, likely driven by the unique job profiles of male part-time workers. At the upper end, men also possess better observables, but this is offset by positive unobserved selectivity among women. The male part-time wage distribution shifted downwards between 2000-2005 and 2012-2017, likely due to worsening wage returns and less specific job profiles. Female part-time wages were more stable over this period, with modest improvements in wage returns but less positive selection in the upper part of the distribution. We observe a convergence of male and female part-time wage distributions, mainly explained by declining differences in wage returns and unobserved selectivity. In conclusion, our study offers compelling evidence for substantial heterogeneity in selectivity patterns across full-time and part-time wage distributions. This phenomenon has not been adequately acknowledged in previous research.

Appendix B

B.1 Multiplier bootstrap

To obtain standard errors and uniform confidence bands of factual and counterfactual distributions as well as of the contributions of individual decomposition factors, we apply a multiplier bootstrap procedure. The validity of this procedure was proven by Chernozhukov et al. (2013, 2025). The multiplier bootstrap is computed by carrying out the following steps for bootstrap replications $b \in 1, ..., B$ and a finite grid of distributional thresholds $y \in \mathcal{Y}$.

(1) Randomly draw bootstrap multipliers $\{\tilde{\omega}_i^b : 1 \leq i \leq N\}$ from the standard exponential distribution. The advantage of draws from the standard exponential is that they can be directly used as weights for computing bootstrapped distribution functions, see below (weights should be nonnegative with mean one). For bootstrapping estimated parameters, a centered version of $\tilde{\omega}_i^b$ is used, i.e.,

$$\omega_i^b = \tilde{\omega}_i^b - \frac{1}{N}\sum_{i=1}^{N} \tilde{\omega}_i^b. \tag{B.1}$$

In order to account for clustering, the same multiplier is drawn for different observations belonging to the same individual i (e.g., Chernozhukov et al., 2020).

(2) The bootstrap estimator of the model parameter at threshold y is obtained as

$$\hat{\theta}_y^b = \hat{\theta}_y + \frac{1}{N}\sum_{i=1}^{N} \omega_i^b \hat{\psi}_i(\hat{\theta}_y), \tag{B.2}$$

where $\hat{\psi}_i(\hat{\theta}_y)$ is the vector of influence functions of the model parameters $\hat{\theta}_y = (\hat{\beta}(y), \hat{\pi}, \hat{\rho}(y))$ for threshold y.

(3) For a functional of interest $\hat{\Delta}_y = \varphi(\hat{\theta}_y, \hat{F}_Z)$ of the model parameters and the sample distribution of observables $\hat{F}_Z(z) = N^{-1}\sum_{i=1}^{N} 1\{Z_i \leq z\}$, compute the bootstrap

realisation of the respective functional's maximal t-statistic as

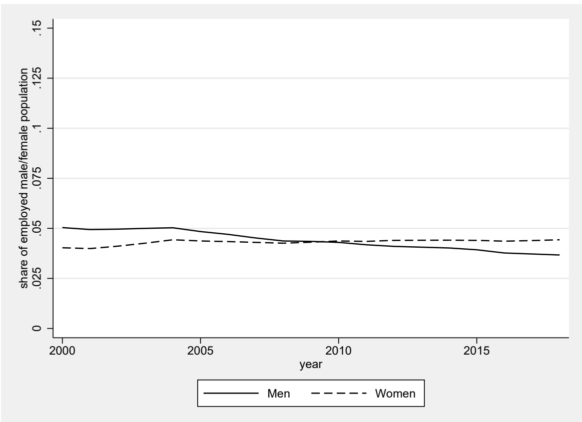
$$t^b_{\mathcal{Y}} = \max_{y \in \mathcal{Y}} \frac{|\hat{\Delta}^b_y - \hat{\Delta}_y|}{SE(\hat{\Delta}_y)}, \tag{B.3}$$

where $\hat{\Delta}^b_y = \varphi(\hat{\theta}^b_y, \hat{F}^b_Z)$ is the bootstrap version of $\hat{\Delta}_y = \varphi(\hat{\theta}_y, \hat{F}_Z)$ based on bootstrap versions of parameters $\hat{\theta}^b_y$ and the bootstrap version of the distribution of observables $\hat{F}^b_Z(z) = N^{-1}\sum_{i=1}^N \tilde{\omega}^b_i 1\{Z_i \leq z\}$. Examples for such functionals are (3.7), (3.10) and (3.15).

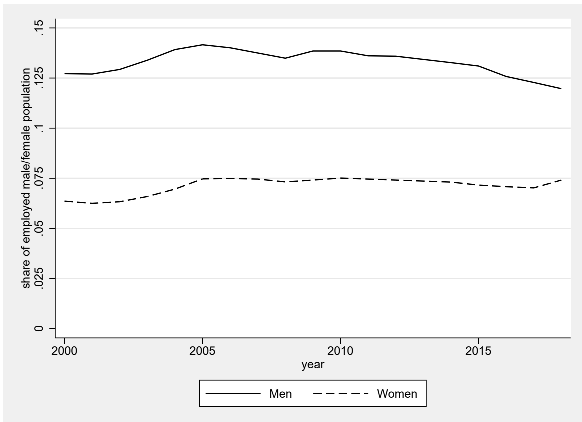
(4) The critical value c_α given a confidence level of $1-\alpha$ can be computed by simulation as the $(1-\alpha)$- quantile of the bootstrap distribution of maximal t-statistics. Uniform confidence bounds are then given by

$$P\left(\hat{\Delta}_y - c_\alpha \cdot SE(\hat{\Delta}_y) \leq \Delta_y \leq \hat{\Delta}_y + c_\alpha \cdot SE(\hat{\Delta}_y) \textit{ for all } y \in \mathcal{Y}\right) \approx 1-\alpha. \tag{B.4}$$

B.2 Self-employment and civil servants



(a) Civil servants



(b) Self-employment

Figure B.1 – Shares of civil servants and self-employed by gender
(Source: Microcensus, Federal Statistical Office)

B.3 Observed characteristics

	Men		Women	
	2000-2005	2012-2017	2000-2005	2012-2017
Selected individuals	$N = 1,458,124$	$N = 1,492,091$	$N = 816,786$	$N = 739,634$
Age	39.503 [10.160]	41.277 [11.092]	38.117 [10.874]	39.438 [11.876]
FT work experience (years)	14.510 [8.352]	17.309 [10.597]	11.207 [7.458]	13.325 [9.615]
FT work experience = 0	0.016 [0.125]	0.020 [0.140]	0.030 [0.169]	0.036 [0.187]
PT work experience (years)	0.176 [0.747]	0.375 [1.241]	0.982 [2.559]	1.645 [3.500]
PT work experience = 0	0.871 [0.335]	0.774 [0.418]	0.683 [0.465]	0.565 [0.496]
Minijob work experience (years)	0.060 [0.296]	0.813 [1.845]	0.133 [0.478]	1.317 [2.355]
Minijob work experience = 0	0.945 [0.296]	0.612 [0.487]	0.878 [0.327]	0.496 [0.500]
Lower/middle sec. school + voc. train.	0.680 [0.466]	0.599 [0.490]	0.639 [0.480]	0.510 [0.500]
Upper sec. school only	0.012 [0.110]	0.022 [0.148]	0.021 [0.144]	0.039 [0.193]
Upper sec. school + voc. train.	0.071 [0.257]	0.112 [0.316]	0.116 [0.320]	0.179 [0.383]
"Fachhochschule" degree	0.054 [0.226]	0.022 [0.146]	0.040 [0.196]	0.024 [0.153]
University degree	0.091 [0.288]	0.173 [0.378]	0.071 [0.257]	0.175 [0.380]
German nationality	0.921 [0.270]	0.912 [0.283]	0.944 [0.230]	0.930 [0.254]
Non-selected individuals	$N = 1,329,506$	$N = 1,282,554$	$N = 1,707,996$	$N = 1,788,235$
Age	40.359 [11.823]	42.704 [11.876]	40.502 [10.784]	42.960 [11.179]
FT work experience (years)	6.180 [7.172]	6.924 [8.279]	5.184 [5.557]	6.498 [6.673]
FT work experience = 0	0.177 [0.382]	0.157 [0.364]	0.194 [0.395]	0.152 [0.360]
PT work experience (years)	0.310 [1.221]	0.742 [2.085]	2.471 [4.625]	3.743 [5.691]
PT work experience = 0	0.824 [0.380]	0.693 [0.461]	0.523 [0.499]	0.377 [0.485]
Minijob work experience (years)	0.133 [0.516]	0.744 [1.716]	0.463 [1.068]	2.079 [3.340]
Minijob work experience = 0	0.886 [0.318]	0.630 [0.483]	0.750 [0.433]	0.431 [0.495]
Lower/middle sec. school + voc. train.	0.546 [0.498]	0.484 [0.500]	0.575 [0.494]	0.515 [0.500]
Upper sec. school only	0.063 [0.244]	0.086 [0.280]	0.045 [0.206]	0.063 [0.243]
Upper sec. school + voc. train.	0.060 [0.237]	0.085 [0.279]	0.077 [0.266]	0.124 [0.330]
"Fachhochschule" degree	0.030 [0.171]	0.025 [0.157]	0.025 [0.155]	0.023 [0.149]
University degree	0.085 [0.279]	0.125 [0.331]	0.066 [0.247]	0.128 [0.334]
German nationality	0.750 [0.433]	0.767 [0.423]	0.857 [0.350]	0.867 [0.340]

Table B.1 – Summary statistics for selection into full-time employment
Note: Standard deviations in parentheses.
(Source: SIAB v7517, own calculations)

	Men		Women	
	2000-2005	2012-2017	2000-2005	2012-2017
Selected individuals	$N = 75,197$	$N = 148,743$	$N = 402,912$	$N = 642,760$
Age	41.387 [13.009]	40.444 [12.083]	42.828 [9.426]	44.324 [10.038]
FT work experience (years)	9.856 [10.031]	9.227 [9.998]	6.541 [5.910]	8.644 [7.048]
FT work experience = 0	0.150 [0.357]	0.141 [0.348]	0.127 [0.331]	0.084 [0.277]
PT work experience (years)	2.779 [3.302]	3.770 [4.124]	7.154 [6.105]	7.924 [6.723]
PT work experience = 0	0.290 [0.454]	0.200 [0.400]	0.010 [0.300]	0.076 [0.265]
Minijob work experience (years)	0.180 [0.530]	1.357 [2.178]	0.222 [0.647]	2.216 [3.263]
Minijob work experience = 0	0.827 [0.378]	0.425 [0.494]	0.830 [0.376]	0.397 [0.489]
Lower/middle sec. school + voc. train.	0.502 [0.500]	0.464 [0.499]	0.691 [0.462]	0.604 [0.489]
Upper sec. school only	0.088 [0.283]	0.078 [0.268]	0.015 [0.120]	0.020 [0.139]
Upper sec. school + voc. train.	0.110 [0.313]	0.133 [0.339]	0.088 [0.283]	0.156 [0.363]
“Fachhochschule” degree	0.055 [0.229]	0.019 [0.137]	0.034 [0.181]	0.023 [0.149]
University degree	0.139 [0.360]	0.200 [0.400]	0.061 [0.240]	0.133 [0.340]
German nationality	0.885 [0.319]	0.858 [0.349]	0.949 [0.221]	0.931 [0.254]
Non-selected individuals	$N = 1,254,309$	$N = 1,133,811$	$N = 1,305,084$	$N = 1,145,475$
Age	40.297 [11.745]	43.000 [11.816]	39.787 [11.072]	42.195 [11.702]
FT work experience (years)	5.959 [6.901]	6.622 [7.977]	4.765 [5.374]	5.294 [6.132]
FT work experience = 0	0.179 [0.383]	0.159 [0.366]	0.215 [0.411]	0.191 [0.393]
PT work experience (years)	0.162 [0.734]	0.345 [1.152]	1.026 [2.763]	1.396 [3.134]
PT work experience = 0	0.857 [0.351]	0.758 [0.429]	0.654 [0.476]	0.546 [0498]
Minijob work experience (years)	0.130 [0.515]	0.664 [1.629]	0.537 [1.157]	2.003 [3.380]
Minijob work experience = 0	0.890 [0.313]	0.657 [0.475]	0.726 [0.446]	0.451 [0.498]
Lower/middle sec. school + voc. train.	0.549 [0.498]	0.486 [0.500]	0.539 [0.499]	0.465 [0.499]
Upper sec. school only	0.062 [0.241]	0.087 [0.281]	0.054 [0.226]	0.087 [0.282]
Upper sec. school + voc. train.	0.057 [0.231]	0.079 [0.270]	0.073 [0.261]	0.106 [0.308]
“Fachhochschule” degree	0.028 [0.166]	0.026 [0.160]	0.022 [0.146]	0.023 [0.150]
University degree	0.082 [0.274]	0.116 [0.320]	0.067 [0.249]	0.126 [0.331]
German nationality	0.742 [0.437]	0.755 [0.430]	0.828 [0.377]	0.830 [0.375]

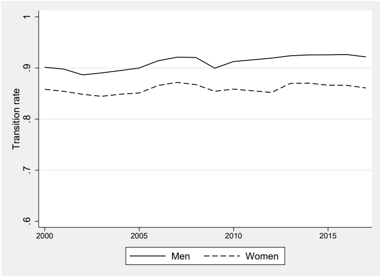
Table B.2 – Summary statistics for selection into part-time employment
Note: Standard deviations in parentheses.
(Source: SIAB v7517, own calculations)

	Outcome	Selection	Sorting
Constant	√	√	√
Year dummies 2001-2017	√		√
Regional (federal state) dummies	√	√	
Age 25-29	√	√	
Age 30-34	√	√	
Age 35-39	√	√	
Age 40-44	√	√	
Age 45-49	√	√	
Age 50-54	√	√	
Age 55-60	√	√	
FT work experience	√	√	
(FT work experience)^2	√	√	
FT work experience x 1(East Germany=1)	√	√	
(FT work experience)^2 x 1(East Germany=1)	√	√	
1(FT work experience=0)	√	√	
PT work experience	√	√	
(PT work experience)^2	√	√	
PT work experience x 1(East Germany=1)	√	√	
(PT work experience)^2 x 1(East Germany=1)	√	√	
1(PT work experience=0)	√	√	
Minijob work experience	√	√	
(Minijob work experience)^2	√	√	
Minijob work experience x 1(East Germany=1)	√	√	
(Minijob work experience)^2 x 1(East Germany=1)	√	√	
1(Minijob work experience=0)	√	√	
Education: lower/middle sec. school + voc. training	√	√	
Education: upper sec. school only	√	√	
Education: upper sec. school + voc. training	√	√	
Education: "Fachhochschule" degree	√	√	
Education: university degree	√	√	
German nationality	√	√	
1(likely student aged 20-25=1)		√	
1(likely student aged 26-30=1)		√	
1(likely student aged 31-35=1)		√	
Full-time share		√	
Part-time share		√	
Minijob share		√	
First difference full-time share		√	
First difference part-time share		√	
First difference minijob share		√	
Transition rate FT → FT		√	
Transition rate FT → PT		√	
Transition rate FT → non-employment		√	
Transition rate PT → FT		√	
Transition rate PT → PT		√	
Transition rate PT → non-employment		√	
Transition rate non-employment → FT		√	
Transition rate non-employment → PT		√	
Transition rate non-employment → non-employment		√	

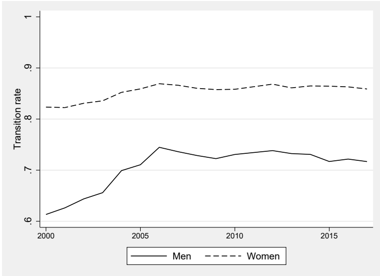
Table B.3 – Covariates used in outcome, selection and sorting equation

Note: Labour market variables used as instruments in selection equation are estimated for cells defined by sex, year, age group as defined above, region (*Raumordnungsregion*) and educational attainment (group 1: secondary school degree only or less; group 2: vocational training; group 3: university/"Fachhochschule" degree)

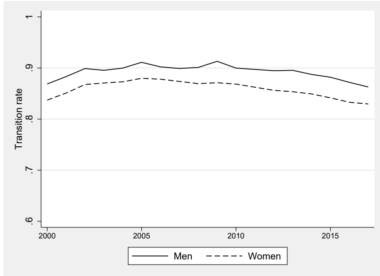
B.4 Time trends in instrumental variables



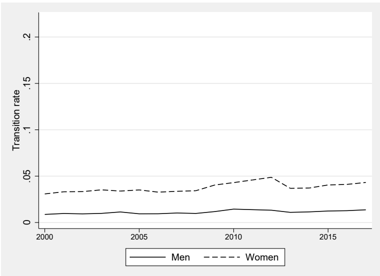
(a) Full-time
→ full-time



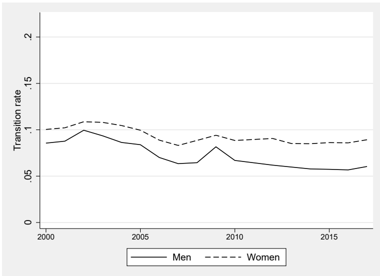
(b) Part-time
→ part-time



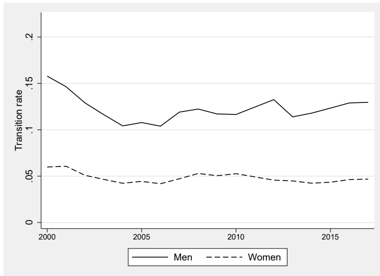
(c) Non-employment
→ non-employment



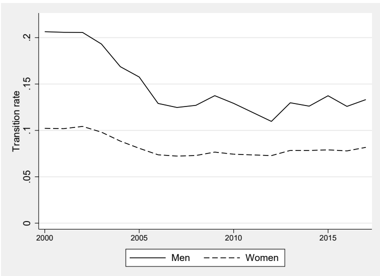
(d) Full-time
→ part-time



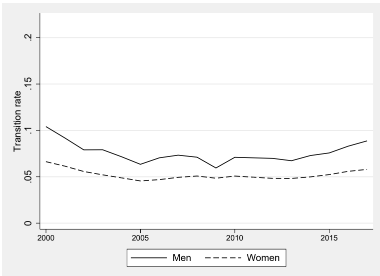
(e) Full-time
→ non-employment



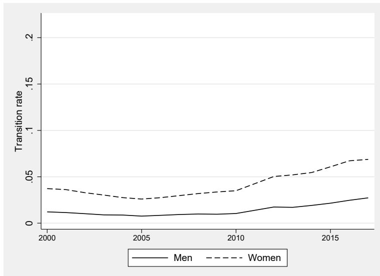
(f) Part-time
→ full-time



(g) Part-time
→ non-employment

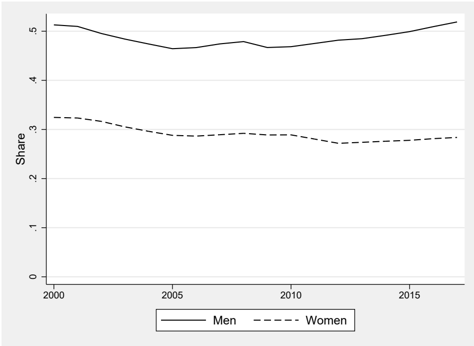


(h) Non-employment
→ full-time

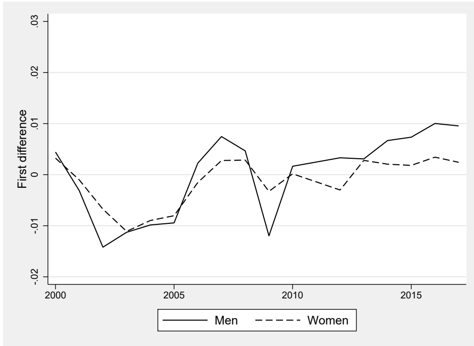


(i) Non-employment
→ part-time

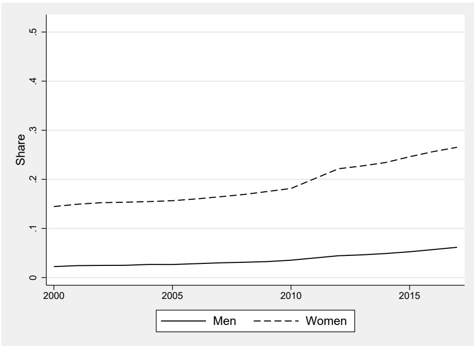
Figure B.2 – Evolution of aggregated transition rates by gender
(Source: SIAB v7517, own calculations)



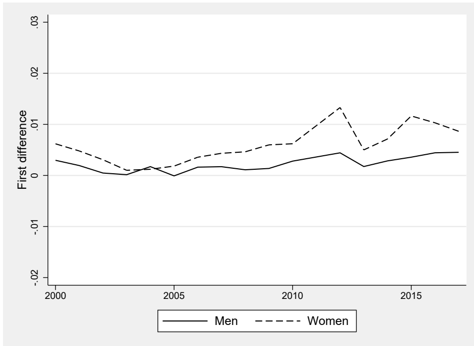
(a) Full-time share



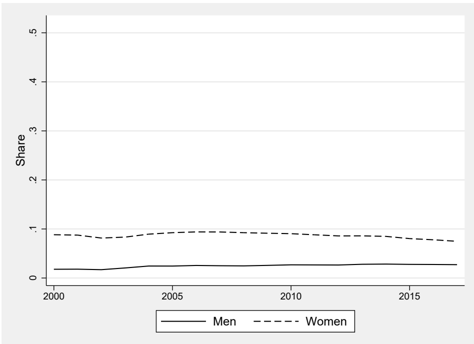
(b) Δ Full-time share



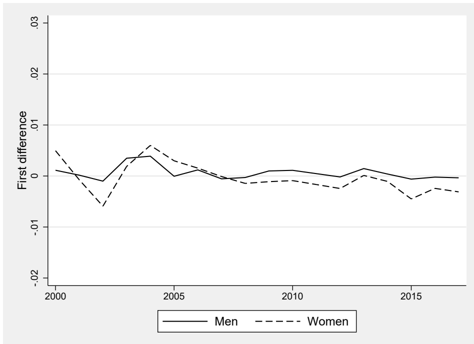
(c) Part-time share



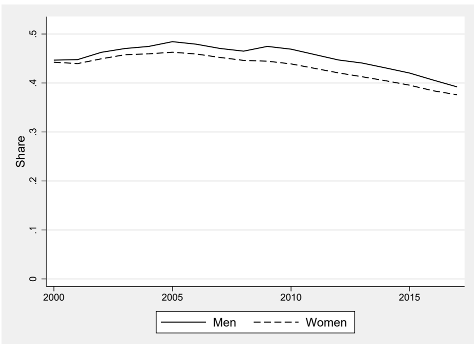
(d) Δ Part-time share



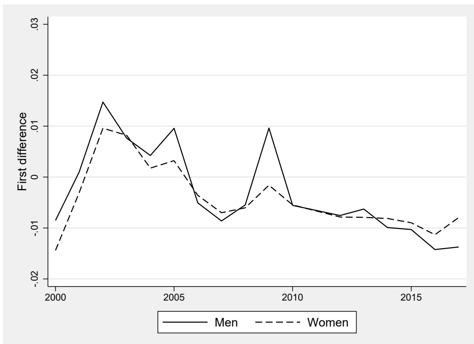
(e) Marginal part-time share



(f) Δ Marginal part-time share



(g) Non-employment share



(h) Δ Non-employment share

Figure B.3 – Evolution of aggregated employment shares
and their first differences by gender
(Source: SIAB v7517, own calculations)

CHAPTER 4

Secular changes in educational attainment and the quality of the highly skilled – Evidence from Germany*

4.1 Introduction

Over the last decades, Germany experienced large changes in the composition of the labour force with regards to educational degrees. As in many other industrialised countries, the shares of university educated individuals have drastically increased for successive cohorts, accompanied by a decreasing share of persons holding a vocational training degree particularly for the younger cohorts (see figure 4.1 for the case of men). Especially after the millennium, the first-year student ratio has been steadily rising, implying a further acceleration of the "academisation" of the German workforce (Wolter and Kerst,

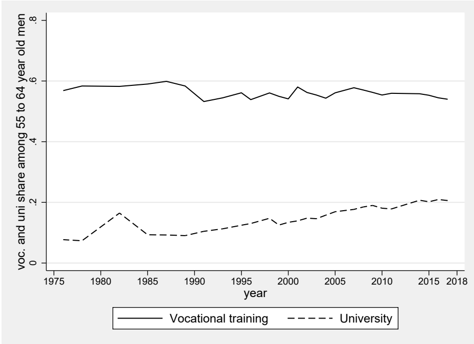
*This chapter is based on: Sturm, Miriam (2025): Secular changes in educational attainment and the quality of the highly skilled – Evidence from Germany, *unpublished manuscript, University of Tübingen.* I thank Martin Biewen and Tobias Brändle, as well as the participants at the International NEPS Conference 2023, the Annual Conference of the International Association for Applied Econometrics (IAAE) 2024, and the Annual Meeting of the Economics of Education Association (AEDE) 2024 for their helpful comments. Maximiliane Sick and Lara Müller provided excellent research assistance.

2015). From an economic perspective, this development has raised at least two concerns. First, Germany's economy is facing an increasing skills shortage in professions requiring vocational training, not least with the so-called "baby boomer" generation approaching retirement age (Deschermeier and Schäfer, 2024; Sauer and Wollmershäuser, 2021). Second, there is a growing debate as to whether the observed educational expansion has led to an increasingly less advantageous selection of young persons into the academic track. This latter point so far lacks empirical evidence due to the inherently unobservable nature of individual ability or quality[1]. Katz and Autor (1999) point out that changes in ability can also be identified more implicitly via the wage structure, arguing that varying educational wage differentials over time may well reflect changes in the average quality within educational groups rather than changes in average wages (given a fixed ability level). Indeed, a number of German studies have documented significant changes in skill premia across cohorts (see Boockmann and Steiner, 2006; Reinhold and Thomsen, 2017, among others), but they have struggled to establish a clear link between these cohort patterns and potential quality shifts due to educational expansion. Doing so is challenging since the evolution of educational shares translates into changes in relative supplies of skill groups, making it difficult to disentangle supply effects from quality effects when relying solely on variation in educational shares over time.

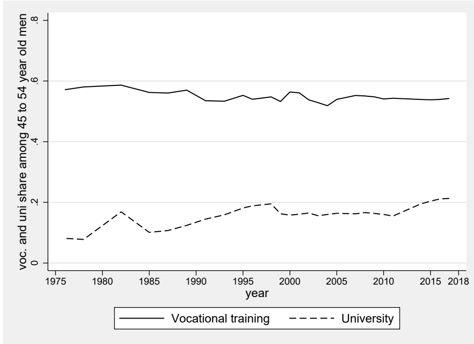
In a similar vein as Carneiro and Lee (2011) for the US, the current study aims to resolve this issue by additionally exploiting regional variation. Using high-quality administrative wage data combined with rich survey data, I compare the full-time wages of highly educated men of the same age working in the same region – thus facing the same relative supplies –, but who received their university entry certificate in different regions. The motivation behind this is that, while they should face the same skill prices on their regional labour market, they might differ in their average ability since the academic track is potentially more selective in some regions compared to others. Selectivity is thereby thought to be measured by the share of a cohort (from the same region of education) that is formally qualified to enrol in the tertiary education system. Wage differences induced by these shares rather than by skill prices are thus assumed to reflect differences linked to the composition of cohorts in terms of their average ability.

This paper contributes to the related literature in several ways. First, it is one among

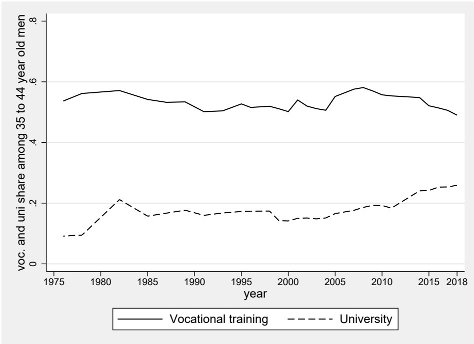
[1]Note that these two terms will be used interchangeably throughout this paper.



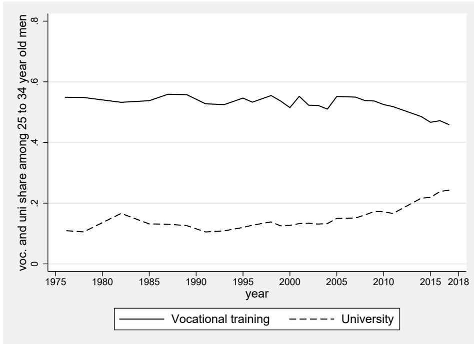
(a) 55- to 64-year old men



(b) 45- to 54-year old men



(c) 35- to 44-year old men



(d) 25- to 34-year old men

Figure 4.1 – Share of men with vocational training/university degree
(Source: Statistical Yearbooks, Federal Statistical Office, own calculations)

few studies based on German data that systematically relate changes in educational attainment over time to changes in wages or skill premia received by different skill groups. Second, the current analysis is – to the best of my knowledge – the first one that explicitly explores the role of changing selection into higher education whereas other studies on Germany focus on the role of an increased supply of college labour for educational returns. In this way, this paper contributes to understanding cohort patterns in skill premia that have been found, but not entirely explained by previous studies. Third, the German case is particularly interesting from an international perspective. One reason for this is that Germany's higher education shares are traditionally low in international comparison, but are starting to catch up with other OECD countries (OECD, 2021), making the pronounced increase in college shares all the more striking. Furthermore, assessing potential quality effects linked to educational expansion is arguably more appropriate in education systems with low educational costs and flat hierarchies of educational institutions. Selection into educational paths should in this case be more tightly linked to ability and less closely connected to financial factors, and changing educational shares should therefore reflect potential ability shifts more accurately. This is in contrast to countries with high educational costs and steep hierarchies such as the US, which the empirical literature has focussed on so far.

The remainder of this chapter is structured as follows. Section 4.2 summarises the literature related to the current study. Section 4.3 gives an overview of the German education system and highlights the features that make the German case particularly well-suited for the given research question and empirical method. Subsequently, section 4.4 introduces the analytical framework. Section 4.5 provides a description of the data used for the analysis, the results of which are presented in section 4.6. Finally, section 4.7 concludes.

4.2 Related literature

There is ample literature on the returns to education in Germany and their evolution over time. In striking contrast to the US where the empirical evidence has documented quite pronounced fluctuations in skill premia (Card and Lemieux, 2001; Katz and Autor, 1999; Katz and Murphy, 1992), educational returns in Germany have been found to be relatively constant over the years (see, e.g., Doepke and Gaetani, 2024; for a comprehensive summary of the empirical evidence until the mid 2000s, see Flossmann and Pohlmeier, 2006). However, an increasing number of studies have shifted their focus away from average cross-sectional returns and towards a more cohort-oriented perspective, unveiling a somewhat more nuanced picture. Based on Mincer-type wage functions including a series of cohort dummies, Boockmann and Steiner (2006) find significant cohort effects in the returns to education, with the latter declining markedly between the cohorts 1945-54 and 1965-74. Gebel and Pfeiffer (2010) confirm these findings for the members of the "baby boomer" generation, who exhibit lower returns to education compared to other cohorts particularly in their early working life.[2] Fitzenberger and Kohn (2006) estimate skill premia for different age groups at different points in time and show that the resulting age profiles have not moved in a parallel fashion over the years, providing further evidence for the existence of cohort effects on top of mere age and time effects. More recently, Reinhold and Thomsen (2017) have found detrimental early labour market outcomes (e.g., slower wage growth) for successive cohorts of young skilled workers entering the labour market around the millennium. The inflow of these cohorts might

[2]Similar cohort effects are found for the UK by Walker and Zhu (2008).

also help to explain the inverted-U shaped pattern in the university premium found by Ordemann and Pfeiffer (2022) and Brüll and Gathmann (2020), who document that the latter has started to decline after 2010.

Evidently, there is a consensus among the aforementioned studies that allowing for cohort effects brings to light heterogeneities that are overlooked when focussing on mere averages. Yet the literature largely leaves open the question what the origin is of the observed cohort patterns in skill premia. Antonczyk et al. (2018) discuss three potential explanations for their existence. First, the labour market conditions upon entry might have long lasting effects on different birth cohorts (Raaum and Røed, 2006; von Wachter, 2020). Second, changes in the demand for and supply of highly skilled workers might cause fluctuations in the university premium over time, particularly in light of skill-biased technological change (see Acemoglu, 2002; Acemoglu and Autor, 2011, among others). A seminal contribution by Card and Lemieux (2001) indeed shows that much of the pronounced rise in the US college premium in the late 20th century can be explained by an accelerated demand for highly skilled labour growing even faster than the supply thereof. Third, changing selection into educational paths may result in differing skill premia between cohorts (Carneiro and Lee, 2011; Juhn et al., 2005).

Both the second and third points are closely linked to the shares of different skill groups over time (and across cohorts). Like many other industrialised countries, Germany has witnessed secular changes in educational attainment over the past decades: While the share of persons with completed vocational training has successively declined in spite of its long history of success in Germany (see, e.g., Riphahn and Zibrowius, 2016), the tertiary education share has risen continuously (Ammermüller and Weber, 2005; Wolter and Kerst, 2015). Although it is rather striking that there seem to be cohort patterns in skill premia and educational attainment alike, there are only few contributions who link the two in a systematic manner. Notable exceptions are Fitzenberger and Kohn (2006) and Glitz and Wissmann (2021) who use CES production frameworks similar to that of Card and Lemieux (2001) to explore the role of relative supply of different skill and age groups for educational returns, thus shedding some light on the second point mentioned by Antonczyk et al. (2018). Glitz and Wissmann (2021), for instance, find that changing relative supplies across cohorts can explain the rise in the vocational training premium (versus low education) fairly well. The evolution of the college premium seems to be somewhat more difficult to retrace, suggesting that there might be something else driving it besides the supply of college labour.

Little attention has so far been paid to the quality of this supply, which connects to the third point mentioned in Antonczyk et al. (2018): changes in selection into tertiary education induced by educational expansion. The rationale behind this is the following: Assuming a time-constant distribution of ability, increasing tertiary education shares might change selection into college in the sense that the "marginal" attendants have less favourable unobservables – henceforth referred to as (academic) ability or quality – than those who would also have participated in tertiary education if college participation rates had been lower.[3] In fact, a few of the studies cited above provide room for speculation about the existence of such quality effects in Germany. The first stage in the control function approach by Gebel and Pfeiffer (2010), for instance, points to a less advantageous selection into tertiary education over time. Likewise, Antonczyk et al. (2018) find negative cohort effects and increasing within-group wage dispersion particularly for the youngest cohorts, potentially reflecting an increased degree of heterogeneity which may be related to a changing distribution of ability within skill groups.

Due to the inherently unobservable nature of ability, it is impossible to (de-)verify this hypothesis in an immediate manner.[4] However – from an economic perspective –, the aforementioned inter-cohort differences in skill premia may provide another way of identifying potential quality effects more implicitly since ability should ultimately be mirrored in wages (Katz and Autor, 1999). To the best of my knowledge, the German literature has hitherto left this possibility unexplored. One of the few studies explicitly linking college premia to a cohort's average ability is that by Juhn et al. (2005) for the US, who model log decade changes in the college premium as a function of log changes in college shares over the same time span. In their setting, college shares are thought to measure a cohort's selectivity or quality, and the authors present tentative evidence for a negative relationship between these shares and college premia. They argue that this is due to a lower average ability with higher enrolment rates, although the estimated effects are not very large and consequently explain only a small part of the college premium

[3]Juhn et al. (2005) present a theoretical framework motivating this hypothesis.

[4]Some studies – most on the US – have relied on proxies such as cognitive test scores in order to explore changing selection into educational tracks (Cunha et al., 2011; Dillon and Veramendi, 2018). However, proxies are not only scarce, but arguably imperfect measures of ability (Heckman and Kautz, 2012).

and its evolution over time.

Measuring ability levels based on educational shares like Juhn et al. (2005) comes with one major challenge: Since educational shares and their evolution over time are tightly linked to relative supplies of different skill groups, it is non-trivial to disentangle the wage effects of supply from the effects of a changing average ability across cohorts. An interesting contribution by Carneiro and Lee (2011) resolves this issue by resorting to another source of variation in educational shares besides variation over time: Using US census data, the authors compare the wages of highly skilled men of the same age who work in the same labour market at the same time (thus facing the same relative supplies), but who were born in different regions with varying shares of cohort members who eventually graduated from college. They find that male college graduates from regions with higher college shares receive lower wages on average than highly educated men from regions with low college shares, which the authors attribute to declining ability levels with higher college shares. The estimated effects are economically large and imply that the increase in the college premium, which was fuelled by a rising demand for college labour (see also Card and Lemieux, 2001), has been significantly dampened by a declining quality of graduates caused by the expansion of tertiary education.

The current paper explores whether similar conclusions can be drawn based on German data. This is an interesting endeavour not only because empirical evidence on the quality effects of educational expansion has yet to be established for Germany, but also because the German education system is especially well-suited for a setting that aims to identify such effects based on regional variation. The next section highlights why this is the case.

4.3 Institutional background

The German education system exhibits several features that contrast substantially with other OECD countries, and that make the German case especially relevant and well-suited to identify potential shifts in the ability of college graduates due to educational expansion. Since the existing empirical evidence on the research question at hand is almost exclusively based on data from the US, this section particularly highlights the differences between the German and US education systems, respectively.

After elementary school (around the age of 10), children can either continue their educational path at lower, middle, or upper secondary school, where the latter is more academically oriented than the former two. The German education system is thus characterised by early tracking, which might be beneficial in terms of specialisation, but can also lead to misallocation of children into the aforementioned tracks and is therefore not uncontroversial (see, e.g., Biewen and Thiele, 2020; Bol and Van de Werfhorst, 2013; Brunello and Checchi, 2007). Selection into secondary schools is based at its core on children's abilities (Cockrill and Scott, 1997; Maaz et al., 2008), although the precise regulations vary over time and also between regions (see last paragraph of this section). Only those who graduate from upper secondary school receive a certificate that qualifies them for higher education, the so-called "Abitur". This certificate generally allows them to enrol either at a university or at a university of applied sciences (called "Fachhochschule"), which are somewhat more practically oriented than the classical universities, but are equally part of the tertiary education system. Unlike colleges in the US, German universities do typically not select their students based on specific aptitude tests which aim to make applicants from different schools more comparable. The decision whether one qualifies for tertiary education is hence made in the secondary school system,[5] which is once more in contrast to the US where selection into higher education is traditionally mostly controlled by colleges themselves (Allmendinger, 1989; Luthra and Flashman, 2017).

Graduates from all three types of secondary schools have the possibility to complete vocational training comprising both classes at vocational school and training at a chosen employer, although employers may have preferences for certain secondary school degrees. The dual vocational training system has a long tradition in Germany and is viewed as one of the pillars of Germany's low youth unemployment rates (Riphahn and Zibrowius, 2016). In the recent past, however, the share of persons with completed vocational training has declined while the tertiary education share has risen continuously (Wolter and Kerst, 2015). These trends are tightly linked to the rising shares of children attending upper secondary school (see also section 4.5) as opposed to lower or middle secondary school. Due to the aforementioned tracking system, it is likely that the drastically rising

[5]In certain cases (i.e., if the person is highly qualified) it is possible to enrol in tertiary education without having graduated from upper secondary school, but these are rare, see Biewen and Thiele (2020).

upper secondary school shares are connected to an increasing number of misallocations to the academic track. This makes countries such as Germany particularly prone to ability shifts induced by educational expansion compared to countries whose education systems are characterised by a lower degree of tracking.

Another central feature of the German education system relates to educational costs. In general, the cost of schooling is fairly low in Germany: Secondary education is typically free of charge, private schools play a negligible role, and universities charge semester fees that are only a small fraction of tuition fees in the US.[6] There are no fees related to vocational schools, and vocational training itself is financed by firms who may pay their apprentices a wage that is, however, typically lower than the wages received by regular employees. In the German case, the largest cost associated with education is the opportunity cost of foregone earnings. Selection into academic versus vocational education (or no professional training at all) should therefore be predominantly based on individual ability and not as much on financial or socioeconomic factors as in other countries. In the US, where there are considerable direct costs of education, changes in tertiary education shares might also be closely linked to fluctuations in tuition, income levels, or the availability of financial aids such as scholarships. Hence, looking at the German case rather than that of the US appears particularly promising if the goal is to link shifts in educational shares to shifts in ability.

Finally – and most important conceptually –, the German education system exhibits a highly federal organisation. Education is generally the responsibility of the 16 federal states, with each of them having their own curricula – although standards are coordinated in the Conference of State Education Ministers ("Kultusministerkonferenz") –, final examinations, and regulations for assigning children to secondary school tracks. The latter may differ in the elementary school grades required to attend upper secondary school, or in the liability to the recommendations given by teachers. In some states, these recommendations are binding whereas in others, parents can decide independently which

[6] There were no general tuition fees at German public universities until 2005, but some federal states introduced fees of up to EUR 500 for some years. These fees are not expected to drive the results of this analysis since the youngest birth cohort contained in the data is that of 1986 (see section 4.5), such that only a small fraction of individuals were affected by these fees. Furthermore, empirical evidence suggests that the introduction of these fees has not had a significant impact on the transition from upper secondary school to university (Baier and Helbig, 2014; Bruckmeier and Wigger, 2014).

track to enrol their children in.[7] While the US secondary school system is characterised by a low degree of standardisation within regions, secondary schooling in Germany is highly standardised within the federal states (Allmendinger, 1989; Henderson et al., 2015), resulting in a much clearer regional distinction. It is this feature that makes the case of Germany particularly appealing to exploit regional variation in order to identify potential quality effects of sorting into higher education. Since secondary schools decide whether an individual qualifies for university or not, the share of persons graduating from school with a university entry certificate in each state (rather than the *college* share in each region of birth as in Carneiro and Lee, 2011) thereby seems a natural measure of selectivity or average ability. The subsequent section describes the analytical framework within which this regional variation is exploited.

4.4 Model

The current study is based on the analytical setting employed by Carneiro and Lee (2011). Let there be two distinct levels of schooling $S = k$, where individuals can either have a college degree ($k = c$) or a vocational training degree ($k = v$). Within each skill type, individual wages may vary between different ages a, years of observation t, regional labour markets r, and the regions where individuals were educated, e. Individual wages for workers of skilly type k can be written as

$$W^k_{iatre} = \Pi^k_{atr} U^k_{i,t-a,e} \quad . \tag{4.1}$$

In this wage equation, Π^k_{atr} denotes the price for k-type skill (e.g., academic skill if $k = c$) among age group a working in region r in year t, and $U^k_{i,t-a,e}$ is the individual-specific endowment with k-type skill for those belonging to cohort $t - a$ who went to secondary school in region e. Log wages are then given in additive form by

$$w^k_{iatre} = \pi^k_{atr} + u^k_{i,t-a,e} \quad , \tag{4.2}$$

with $w^k_{iatre} = \log W^k_{iatre}$, $\pi^k_{atr} = \log \Pi^k_{atr}$, and $u^k_{i,t-a,e} = \log U^k_{i,t-a,e}$.

[7]Teacher recommendations are controversially discussed in the literature: While Grewenig (2021) finds positive effects of binding recommendations on student achievement, Batruch et al. (2023) argue that teacher recommendations tend to be biased based on the socioeconomic background of students.

In order to derive expected wages for certain groups of individuals, define indicators S_i^k for having educational degree k, A_{iat} for belonging to age group a in year t, and M_{itre} for having been educated in region e and working in region r in year t. Further, denote the expected log wage for someone with qualification k belonging to age group a in period t who was educated in region e and works in region r by $\omega_{atre}^k = E[w_{iatre}^k | S_i^k = 1, A_{iat} = 1, M_{itre} = 1]$, and analogously the expected (log) endowment of k-type skill by $\nu_{atre}^k = E[u_{i,t-a,e}^k | S_i^k = 1, A_{iat} = 1, M_{itre} = 1]$. Together with equation (4.2), this yields

$$\omega_{atre}^k = \pi_{atr}^k + \nu_{atre}^k \quad , \tag{4.3}$$

such that expected log wages are given for cells defined by (k, a, t, r, e).

Equation (4.3) has the following important implications. First, expected log wages are the sum of the log price π_{atr}^k received by k-type skill workers of a similar age a working on the same labour market r at time t, and the average endowment of these workers with k-type skill, ν_{atre}^k. The latter can be thought of as measuring the average quality of workers in the respective cell. Second, while skill *prices* are identical among individuals belonging to the same (k, a, t, r) cell, average skill *endowment* may additionally depend on the region where individuals were educated, i.e., it may vary with (k, a, t, r, e). This implies that exploiting regional variation makes it possible to disentangle supply effects (which are absorbed by π_{atr}^k) from quality effects (reflected by ν_{atre}^k). Put differently, differences in the average quality of workers can be recovered by comparing persons who were educated in different regions, but who otherwise face the same skill prices π_{atr}^k due to them being of the same age and working in the same labour market region at the same time.

Part of the difference in average quality between regions of education e might be the result of varying shares of persons holding a university entry certificate ("Abitur") across regions e (and cohorts $t - a$). The rationale behind this is straightforward: Assuming that the distribution of ability is the same across regions and over time, the "Abitur" shares of the respective cohorts should be approximately equal. If they are not, this might indicate, for instance, that upper secondary schools are more selective in some federal states than in others, perhaps due to differing regulations regarding the transition from elementary to secondary school in the tracking system. As discussed in section 4.3 (and as will be shown in section 4.5), the federal organisation of the German education system has the potential to – and does – create considerable variation in "Abitur" shares. The latter shall therefore serve as the central quality measure in the remainder of this study.

It is important to think about how to model average skill endowment (ν^k_{atre}) and prices for k-type skill (π^k_{atr}). As mentioned before, prices only vary at the level of (a, t, r) within each skill type k. To avoid any functional-form restrictions, they are modelled as year–age–region-of-work fixed effects γ^k_{atr} by including dummies for each (k, a, t, r) cell. As to average k-type skill endowment within an (a, t, r, e) cell, it seems likely that there are also other factors besides the share of cohort members eligible for college that result in differences in ν^k_{atre} across regions e. Other characteristics of the regions in which individuals were educated – and where they most likely grew up – might also affect skill endowment within a cell. To name one example, individuals growing up in structurally weak regions might not have the same opportunities as their peers from other regions in terms of, e.g., early labour market experience or parental income, such that average skill endowment in these cells might be lower compared to cells corresponding to other regions of education e. Furthermore, vocational and academic education might be valued and promoted differently in different regions of education, such that a region might be relatively well-endowed with v-type skill while simultaneously exhibiting relatively low endowment with c-type skill compared to other regions. The specification therefore potentially also includes region-of-education fixed effects as well as their interactions with year of observation t, age a, and region of work r, respectively. Finally, there might be selective migration: As an example, high-ability workers may tend to move to certain regions r, or they may tend to leave certain regions e. To account for possible wage differences induced by such migration flows from region e to labour market r, a function $\lambda^k(P^k_{M,atre}, P^k_{M,atrr})$ of the migration probabilities $P^k_{M,atre}$ and staying probabilities $P^k_{M,atrr}$ in each cell may also explain part of ν^k_{atre}.[8]

Considering all these potential confounders yields the following model. Let $P_{t-a,e}$ denote

[8]This means that for each cell (a, t, r, e) and skill group k, $P^k_{M,atre}$ is estimated as the fraction of k-skilled, a-aged individuals from region e working in region r in year t, i.e., it reflects how "popular" region r is as a destination for individuals from region e. Similarly, $P^k_{M,atrr}$ reflects how likely it is that individuals educated in region e stay to work in this region later on, i.e., $e = r$.

the "Abitur" share of cohort $t-a$ from region e, and $\phi^k(P_{t-a,e})$ some function thereof. Using that $\pi^k_{atr} = \gamma^k_{atr}$ and $\nu^k_{atre} = \gamma^k_{ae} + \gamma^k_{te} + \gamma^k_{re} + \phi^k(P_{t-a,e}) + \lambda^k(P^k_{M,atre}, P^k_{M,atrr})$, equation (4.3) becomes

$$\omega^k_{atre} = \gamma^k_{atr} + \gamma^k_{ae} + \gamma^k_{te} + \gamma^k_{re} + \phi^k(P_{t-a,e}) + \lambda^k(P^k_{M,atre}, P^k_{M,atrr}) \quad , \tag{4.4}$$

with γ^k_{ae}, γ^k_{te} and γ^k_{re} being full fixed effects by region of education and age, year, and region of work, respectively. As discussed above, these fixed effects aim to control for a variety of factors affecting average quality within cells. After controlling for these effects and for skill prices γ^k_{atr}, the remaining wage differences between cells must be due to differences in quality due to the selectivity of upper secondary schooling. In practice, $\phi^k(P_{t-a,e})$ will be a linear function of the odds ratio of the share of cohort members from the same region e holding a university entry certificate, i.e., $\phi^k(P_{t-a,e}) = \beta^k \frac{P_{t-a,e}}{1-P_{t-a,e}}$. The odds ratio is a convenient choice since it provides more variation than the "Abitur" share itself (which by definition only varies between 0 and 1, while the odds ratio may take on values between 0 and ∞). Estimates of migration and staying probabilities $(P^k_{M,atre}, P^k_{M,atrr})$ are (if considered at all) included as second-degree polynomials.

Equation (4.4) is estimated separately for average wages received by persons who have completed college ($k = c$) or vocational training ($k = v$), respectively, rather than using the college *premium* as the dependent variable. The reason for this is that varying college premia might not be exclusively caused by changes in the quality of college graduates, but possibly also by changes in the quality of those with completed vocational training. It seems possible that the quality of the latter group is also affected by rising "Abitur" shares since these imply that an increasing share of young persons have the alternative to pursue a college degree, which may alter the pool of applicants for vocational training. If such spillover effects exist, these would blur the effect of sorting into higher education when focussing on the college premium. Looking at the two groups separately makes it possible to assess to what extent educational expansion is linked to selection into the tertiary education system *and* into vocational training, yielding a more holistic picture.

4.5 Data

4.5.1 Data sources

Several data sources are combined to examine the relationship under question for the case of Germany. First and foremost, the current analysis hinges on the use of high-quality wage data, which are taken from the Sample of Integrated Labour Market Biographies (SIAB). The SIAB is a 2% random sample drawn from the Integrated Employment Biographies (IEB), which contain exact to-the-day information on all employees covered by social security records between 1975 and 2017. The records of employees in the East German states are available in the data starting in 1992. All in all, the SIAB records (within the given time frame) the entire employment histories of approximately 1.8 million individuals. On the one hand, its administrative nature makes it perfectly suited for applications relying on high-quality data on the daily wage received by employees. On the downside, the IEB only record employment subject to social security contributions, meaning that public servants and self-employed individuals are not contained in it. A related problem is that wages are only reported up to the social security contributions ceiling. Right-censored wages are therefore imputed following Gartner (2005), as is common practice in the related literature (see appendix C.1 for details).

In spite of its aforementioned advantages, the SIAB provides very little information on the individuals whose administrative records it is based upon. Unlike the given wage and employment information which is precisely measured, the SIAB only contains very basic and partially incomplete information on individuals' education. The current study therefore makes use of the so-called NEPS-SC6-ADIAB database, a project conducted jointly by the IAB and the Leibniz Institute for Educational Trajectories (LifBi). The latter is the provider of the National Educational Panel Study (NEPS, for details see NEPS Network, 2022), a unique survey providing detailed information on the entire educational paths of roughly 17,000 survey participants. The NEPS-SC6-ADIAB project links the survey respondents (conditional on their agreement[9]) to their administrative records, making it possible to observe the entire employment biographies of the survey participants. Thus, the resulting dataset combines reliable wage data with rich information on the educational background of the observed individuals.

[9]Roughly 93% of the survey respondents agreed to the linkage, see Bachbauer et al. (2021).

Among other things, the survey participants report in which year and federal state they acquired each of their educational degrees. As laid out in the previous section, this is precisely the sort of information needed to disentangle supply effects from quality effects based on regional variation. The NEPS data are arguably more suitable in the given context than that used by Carneiro and Lee (2011) whose census data only report the region of birth, which they need to assume to also be the region where individuals were educated. This is potentially problematic especially in case of the US where (1) due to the steep hierarchy of higher educational institutions, there's sizeable migration across regions especially of the most qualified persons who may wish to attend a prestigious college elsewhere (Faggian and Franklin, 2014), and at the same time (2) colleges are the institutions that ultimately decide who can enrol, as opposed to Germany where the decision whether someone qualifies for college is primarily made in secondary schools (see section 4.3). Thus, the assumption made by Carneiro and Lee (2011) that the region of birth is also the region where individuals were selected for college might be inappropriate especially for highly able persons. The detailed information made available in the NEPS prevents related complications that may interfere with the results.

As mentioned before, the central measure of selectivity – and hence of a cohort's average (academic) ability – is the share of young persons graduating with a university entry certificate in a given year and federal state. Since the sample size is limited in the NEPS, these "Abitur" shares are instead computed based on data gathered from the Statistical Yearbooks. The latter are rich collections of economic, demographic and other statistics provided by the Federal Statistical Office of Germany, including the exact number of individuals who successfully graduated from upper secondary school in each year and federal state. Together with the size of the relevant age group[10]– which is also reported in the Statistical Yearbooks –, these numbers can readily be used to compute the "Abitur" share of a given cohort defined by graduation year and region. The gathered data reach back until the mid 1950s, ensuring that each person in the NEPS-SC6-ADIAB can be assigned her or his "Abitur" share. Combining these data sources results in a dataset that is tailored specifically for the given analytical setting.

The analysis focusses on full-time working men aged 25 to 64 years due to the high

[10]The Statistical Yearbooks report the number of state inhabitants aged 18 to under 21 years. The "Abitur" share is computed by dividing the number of graduates by one third of the size of this age group.

part-time share of women and the absence of information on hours worked in the administrative data. For individuals with several employment spells per year, the wage information from the longest employment spell of a given year is determined and included into the analysis.[11] In the rare case of parallel full-time employment spells, I only use the one with the highest reported wage. Furthermore, only wages observed between 1999 and 2017 are considered since the NEPS only contains birth cohorts 1944 to 1986 and the analysis groups individuals into 5-year age groups (see below), such that the chosen years are the only ones that are relatively balanced in composition with regards to age.

4.5.2 Descriptive evidence

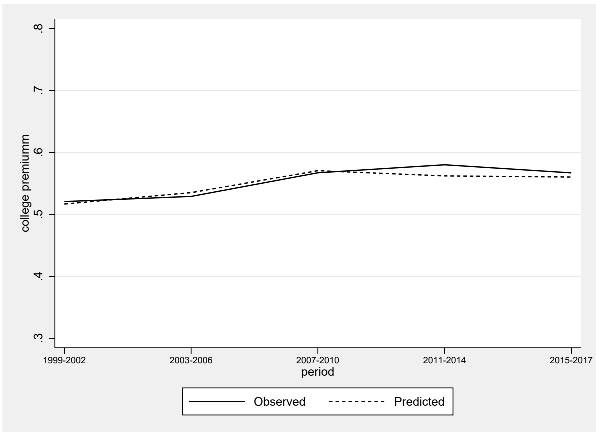


Figure 4.2 – Aggregated college premium using observed and predicted average wages (Source: NEPS-SC6-ADIAB, own calculations)

The solid line in figure 4.2 depicts the evolution of the college premium – estimated as the simple log difference between average college and vocational wages – for the given sample.[12] It appears that the German college premium is not very dynamic, exhibiting only a slight upward trend from about 53 to 58% over a time span of almost 20 years and recently even a slight decline (see also Brüll and Gathmann, 2020; Ordemann and Pfeiffer, 2022). As highlighted in section 4.2, this is in stark contrast to the US college premium, which has evolved much more dynamically (Doepke and Gaetani, 2024).

[11]The analysis is generally based on real wages inflated based on the Consumer Price Index (2017=100) provided by the Federal Statistical Office.

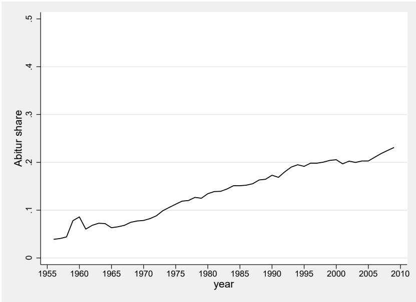
[12]Note that the dashed line is a predicted premium obtained from a version of equation (4.4) and that the premia are given for subperiods comprising 4 years (see reason below).

The lack of movements in the college premium also makes it difficult to surmise in what way the latter might be related to changes in educational shares, which have evolved much more dynamically: As it turns out, there is considerable variation in “Abitur” shares – over time, but also between different federal states. Figure 4.3 exemplarily depicts their evolution for four states.[13] Panels (a) to (c) reveal the general trend of an increase in the share of young persons eligible for college. However, this trend accelerates in a much more pronounced manner in the states of North Rhine-Westphalia and Bremen compared to Bavaria, particularly in the 1970s. Over the time span between 1955 and 2010, “Abitur” shares rose from below 10% in 1955 to 23% in Bavaria, 32% in North Rhine-Westphalia, and even 38% in the state of Bremen. Similar upward trends – albeit at a lower level – can be observed for the first-year student ratios in the same states shown in figure C.2 in appendix C. This confirms that the federal organisation of the German education system produces a significant amount of regional differences, which makes it promising to pursue the approach outlined in the previous section. Interestingly, the periods in which “Abitur” shares (and first-year student shares) show the most pronounced increase are also the ones in which college dropout rates have risen most markedly (see figure C.3 in appendix C). This is especially true of the aforementioned surge in the 1970s, and could be a first hint at a decline in the average ability of those admitted into the tertiary education system as relatively more persons formally qualify for it.

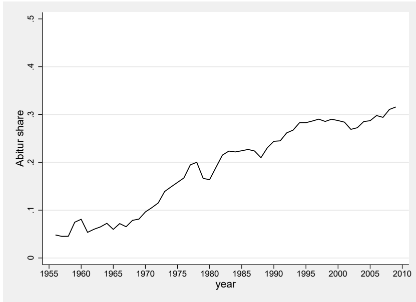
Panel (d) depicting the evolution of the “Abitur” shares for the East German state of Brandenburg highlights a particular source of variation coming from individuals who were educated in the German Democratic Republic (GDR). In contrast to the Federal Republic of Germany where “Abitur” shares already rose rather sharply before the re-unification, they remain just under 10% in Brandenburg and the other regions of the GDR[14] where admission into the academic track was especially restrictive. Since it is not clear whether admission was solely based on ability or possibly also on the family's party affiliation or political activities in general, the empirical part will consider different samples, where one variation will omit cells in which the GDR is the region of education, and a third variation additionally excludes East German states as the region

[13] The evolution of “Abitur” shares in the remaining 12 states is depicted in figure C.1 in appendix C.

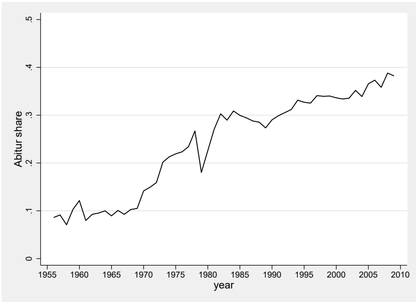
[14] Note that population numbers and the number of upper secondary school graduates are only given for the whole GDR rather than for different regions, such that the shares reported before the reunification are the same across all states in the GDR.



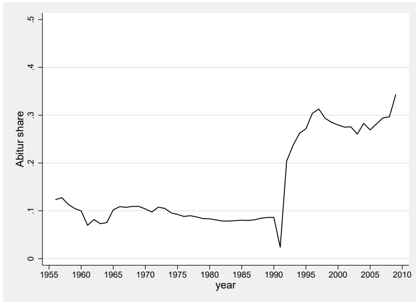
(a) Bavaria



(b) North Rhine-Westphalia



(c) Bremen



(d) Brandenburg

Figure 4.3 – Shares of young persons graduating with "Abitur"
(Source: Statistical Yearbooks, Federal Statistical Office, own calculations)

of work.

In an ideal setting, one would estimate the model specified in equation (4.4) based on fine cells defined by 16 regions of education (17 if those educated in the GDR are included as a separate group) and 16 regions of work to fully capture the heterogeneity between these federal states. However, using the detailed survey data from the NEPS unfortunately comes at the cost of a limited number of observations. Distinguishing all 16 regions would produce an abundance of empty cells and hardly any cells with more than a handful of observations. This limitation necessitates an aggregation of different federal states into regions that are comparable in terms of their "Abitur" shares across time, and that are also similar in terms of other factors that might affect the labour market situation.

To substantiate the choice of aggregated regions, figure C.4 in appendix C exemplarily shows in which federal states "Abitur" shares were relatively high or low in the years 1975, 1985 and 1995, with darker shades indicating higher shares. What catches the eye is that the three city-states Berlin (West Berlin before the reunification), Hamburg and Bremen exhibit the highest shares of young persons graduating with a general qualification for

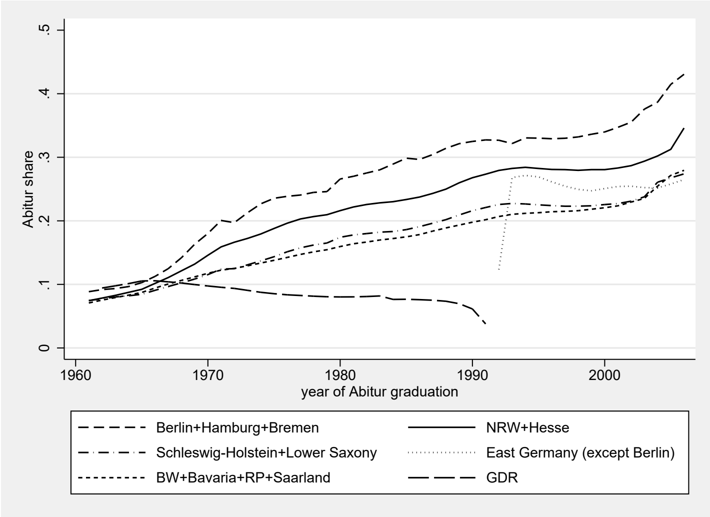


Figure 4.4 – Aggregated shares of young persons graduating with "Abitur"
(Source: Federal Statistical Office, Statistical Yearbooks, own calculations)

higher education. Since they should also be rather comparable in other respects (all are cities, located in the northern part of Germany and entirely surrounded by larger states), it seems sensible to aggregate these three states. The states of North Rhine-Westphalia and Hesse often have the next highest "Abitur" shares. Given that the two states are also geographically close and characterised by both rural and metropolitan areas (the Ruhr area in case of North Rhine-Westphalia and Frankfurt in case of Hesse), they are also aggregated. The other aggregates are Schleswig-Holstein and Lower Saxony (both located in the North), and Baden-Wuerttemberg, Bavaria, Rhineland-Palatinate and Saarland (located in the South), where the states in these aggregates have relatively low "Abitur" shares. Finally, the last group of federal states are those in the East that once belonged to the GDR, and that exhibit the lowest "Abitur" shares before the reunification and considerably higher shares afterwards.

The resulting cells are thus built based on 5 regions of work and 6 regions of education, where the latter include the GDR. An overview of these regions is provided in table C.1 in appendix C. Besides aggregating over federal states, workers of a similar age are also grouped together into 5-year age groups (25-29, 30-34, ..., 60-64 years of age). Further, to increase cell sizes for the sake of a more robust estimation of average wages ω^k_{atre}, the final cells always comprise four neighbouring years. Given that each cell must be attributed one unique "Abitur" share, this further aggregation with respect to age a and year t necessitates the use of an average "Abitur" share over time rather than the

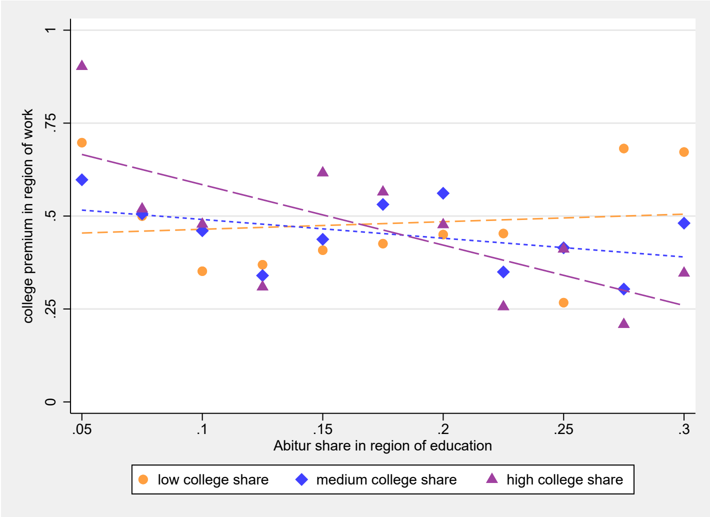


Figure 4.5 – College premium in region of work as a function of "Abitur" share in region of education

(Source: NEPS-SC6-ADIAB, own calculations)

Note: Categories low, medium and high are based on the terciles of the distribution of college labour shares.

exact year-by-year shares presented above. The resulting aggregated shares are depicted in figure 4.4 and clearly show that there is still considerable variation in the shares of young persons graduating with "Abitur" over time and between the more broadly defined regions. As noted in the previous sections, this regional variation is central to the idea of the chosen approach.

If "Abitur" shares are telling of the average ability of college graduates educated in the different regions, these differences should ultimately also be mirrored in their wages. But to what extent is this the case? Figure 4.5 gives a first glimpse at the relationship the current analysis aims to measure. The points shown in this graph represent college premia (college vs. vocational wages) estimated for cells defined by year, age, region of work, and the region where individuals went to secondary school. These premia are computed separately for different types of labour markets with varying relative supplies of college labour, which aims to control – to a certain extent – for supply effects. The resulting college premia are then plotted against the corresponding "Abitur" share in the region where individuals received their qualification for higher education. Figure 4.5 offers two main insights. First, there tends to be a negative relationship between college premia and "Abitur" shares. Following the argument that motivates the current study, this is potentially due to a declining average ability with increasingly higher shares of persons formally qualified to enrol in the tertiary education system. Second, the strength of this relationship seems to depend on the relative supply of college labour. While there is no clear pattern for labour markets with a relatively low share of highly skilled persons (orange), it becomes negative for those with a medium supply (blue) and is steepest for labour markets with the highest shares of college labour (purple). The next section reveals whether this negative association is still there in an analytical setting that models demand and supply effects more flexibly as lined out in section 4.4.

4.6 Empirical results

4.6.1 Regression results

Table 4.1 presents the results obtained from variations of equation (4.4). These variations differ in the minimum cell size imposed[15] as well as in the exact set of controls included. Each specification is carried out (a) including every cell, (b) excluding cells in which the GDR is the region of education, and (c) additionally excluding East Germany as the region of work. As noted in section 4.4, the dependent variables are not college *premia*, but average log college and vocational wages in each cell. Due to the multitude of fixed effects included in the model, table 4.1 only reports the estimated coefficient and standard error on the covariate the current analysis focusses on, which is the odds ratio of "Abitur" shares. Recall that this variable is intended to measure the composition of a cohort with regards to its average ability. Hence, the reported coefficient estimates aim to measure the wage effects of educational expansion that can be attributed to differences in average ability between cells.

All of the uneven columns report the results for college wages. The main takeaway from the given estimates is that they all point to a negative relationship between "Abitur" odds ratios and college wages, implying that individuals from regions (and cohorts) with higher "Abitur" shares tend to receive lower wages on average. This is consistent with the hypothesis that educational expansion draws increasingly lower abilities into tertiary education, and aligns with the results obtained by Carneiro and Lee (2011) for the US. Irrespective of the set of fixed effects included, the estimate of interest remains rather

[15]Tables C.2 and C.3 in appendix C show the composition of the samples used in the regressions, displaying the fraction of cells belonging to each age group a, region of work r, region of education e, and time period t.

robust in size, ranging from about -1.2 to -1.0 when including cells with the GDR as the region of education and from -1.0 to -0.7 when exclusively considering cells from West Germany. But what do these numbers tell us? It turns out that taking even the most conservative estimate obtained from specification (1c) yields an effect that is non-negligible in magnitude. To see this, recall that the presumed quality effect is measured by a linear function of the odds ratio of "Abitur" shares, i.e., by $\phi^k(P_{t-a,e}) = \beta^k \frac{P_{t-a,e}}{1-P_{t-a,e}}$. For the sake of simplicity, assume that the "Abitur" share within a cell defined by $(t-a, e)$ doubles from 10 to 20%, which roughly corresponds to the increase observed in Bavaria over the whole time span since 1970 and to the striking expansion in the state of North Rhine-Westphalia in the 1970s alone (see figure 4.3). A rise in "Abitur" shares of this magnitude would imply an increase in the odds ratio by about 0.139, which – combined with $\hat{\beta}^c = -0.692$ – yields a detrimental effect on college wages by $-0.692 \cdot 0.139 = -0.096$, i.e., -9.6% that is attributable to a decline in the average ability of college graduates. Relying on the larger (and statistically significant) coefficient from specification (1a) based on *all* cells would imply that this same rise in "Abitur" shares is accompanied by a quality-related decline in college wages by about -15.8%, which is quite considerable.

While unanimous in their sign, the coefficient estimates are not always statistically significant. By way of example, the presumed quality effect is – as just noted – very large and highly significant when including cells in which the GDR is the region of education, but it becomes substantially smaller and is no longer significant when these cells are omitted (see specifications (1b) and (1c)). As expected, the estimates within the same column are considerably more robust to omitting the GDR when the minimum cell size is increased. This becomes most evident by comparing the first to the third column, which contains the same set of controls, but requires cells to contain at least 10 (instead of 5) observations.[16] In the third column, the coefficient estimate on the "Abitur" odds ratio only drops from -1.021 in specification (3a) to -0.920 in specifications (3b) and (c) while also remaining significant on a level of 10%. Similarly modest drops in the estimated parameter of interest are visible in columns (5) and (7), suggesting that the estimated effects are not solely driven by the cells belonging to the former GDR.

[16]Note that the samples based on cells with at least 5 and at least 10 observations, respectively, are very similar in composition (see tables C.2 and C.3 in appendix C).

	(1) $\ln(\text{wage}^c)$	(2) $\ln(\text{wage}^v)$	(3) $\ln(\text{wage}^c)$	(4) $\ln(\text{wage}^v)$	(5) $\ln(\text{wage}^c)$	(6) $\ln(\text{wage}^v)$	(7) $\ln(\text{wage}^c)$	(8) $\ln(\text{wage}^v)$
(a) Including e=GDR and r=East Germany								
Odds ratio "Abitur"	-1.140***	-0.525**	-1.021*	-0.650***	-1.212**	-0.574**	-1.266	-0.301
	[0.407]	[0.207]	[0.533]	[0.230]	[0.598]	[0.266]	[1.753]	[0.551]
R^2	0.892	0.945	0.931	0.954	0.926	0.940	0.943	0.961
N	576	624	380	498	380	498	380	498
(b) Excluding e=GDR, including r=East Germany								
Odds ratio "Abitur"	-0.696	0.003	-0.920*	-0.008	-0.914	-0.333	-1.061	-0.386
	[0.501]	[0.369]	[0.546]	[0.407]	[0.739]	[0.479]	[1.998]	[0.607]
R^2	0.903	0.913	0.921	0.926	0.913	0.917	0.933	0.942
N	450	491	296	378	296	378	296	378
(c) Excluding e=GDR and r=East Germany								
Odds ratio "Abitur"	-0.692	0.003	-0.920*	-0.008	-0.914	-0.333	-1.061	-0.386
	[0.490]	[0.362]	[0.538]	[0.402]	[0.727]	[0.473]	[1.967]	[0.599]
R^2	0.893	0.892	0.910	0.905	0.902	0.893	0.924	0.925
N	424	472	287	368	287	368	287	368
Min. cell size	5	5	10	10	10	10	10	10
a-t-r fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
e-a fixed effects	YES	YES	YES	YES	YES	YES	YES	YES
e-r fixed effects	YES	YES	YES	YES	NO	NO	YES	YES
e-t fixed effects	NO	NO	NO	NO	NO	NO	YES	YES
Migration rates	NO	NO	NO	NO	YES	YES	YES	YES

Table 4.1 – Regression results

(Source: NEPS-SC6-ADIAB, own calculations)

Note: Regressions of average college and vocational wages within (a, t, r, e)-cells (age, year, region of work, region of education) on "Abitur" odds ratio within $(t - a, e)$-cells (year–age, region of education). Observations (i.e., cells) are weighted by the inverse sampling variance of log college/vocational wages within cells. Heteroskedasticity-robust standard errors in parentheses. *: significant at 10 % level; **: significant at 5 % level; ***: significant at 1 % level.

Besides these differences between the rows of the same column, there are also some differences between columns. As previously mentioned, the estimated quality effect is overall rather robust to variations in the set of controls, irrespective of whether only $(e-a)$ and $(e-r)$ fixed effects are included (columns (1) and (3)), whether $(e-r)$ fixed effects are replaced by migration rates (column (5)), or whether one additionally includes $(e-t)$ fixed effects as well as migration rates (column (7)). When including the GDR into the regressions, the parameter of interest is also statistically significant in almost all specifications. The only exception is (7a), where the introduction of $(e-t)$ fixed effects in addition to the fixed effects in the previous columns appears to introduce some multicollinearity, thus inflating the standard errors while the estimate itself remains similar to its counterparts in the other columns. The picture is somewhat different when excluding the GDR as the region of education (and East Germany as the region of work), where only specifications (3b) and (3c) report a statistically significant relationship between the "Abitur" odds ratio and college wages, and that only on a 10% level of significance. This is not surprising in light of the fact that omitting cells with the GDR as the region of education results in a significantly lower number of observations, and that these cells – with their unusually low and stable "Abitur" shares (recall figure 4.3) – furthermore contribute a considerable amount of variation to the specifications in the first row that is missing from the remaining two. The direction of the estimated effect, however, is unambiguous across all specifications.

All in all, table 4.1 tends to provide evidence of declining college wages with higher shares of persons considered as qualified for enrolling in college, albeit with varying degrees of significance. This is consistent with comparable empirical evidence on the US established thus far (Carneiro and Lee, 2011; Juhn et al., 2005) and at least suggestive of the hypothesis that educational expansion draws individuals located towards increasingly lower percentiles of the ability distribution into higher education. Due to the limited sample size, the magnitude of the estimated effects should nevertheless be treated with caution.

It is further important to mention that the source of the presumed quality effect is not obvious. As also noted by Carneiro and Lee (2011), a declining "ex-ante quality" of college participants in the sense that they already differ in terms of their inherent ability before they work their way through the education system may only be part of the story. An alternative explanation is that unless financial and personnel resources are adequately raised, the amount of resources available per student declines the more young persons attend upper secondary schools (and subsequently colleges). This latter point may impair the quality of education itself, resulting in yet another detrimental effect on the quality of college graduates. The coefficients presented in table 4.1 can therefore be thought of as measuring the gross quality effect of educational expansion. Since the two distinct sources would call for crucially different policy implications[17], table C.4 in appendix C makes an attempt to separate them by additionally controlling in the specifications for college wages from table 4.1 for the financial resources of German educational institutions. Based once more on information from the Statistical Yearbooks, the regressions presented in table C.4 control for real public expenditures for universities per first-year student and for upper secondary schools per graduate, respectively, where expenditures are given by region of education e and cohort $(t - a)$ (i.e., expenditures are given at the same level as "Abitur" shares).[18] As it turns out, the estimates on the "Abitur" odds ratio remain rather robust to the inclusion of university expenditures, but they become somewhat smaller in magnitude when controlling for upper secondary school expenditures. I conjecture that this is suggestive of the *gross* quality effect in table 4.1 being driven both by a decline in "ex-ante ability" and by an impaired quality of (upper secondary) education due to educational expansion. While this seems conclusive, the estimates in table C.4 should once more be taken with a grain of salt due to the low sample size, which is impeded by the lack of expenditure data for the former GDR (which is why there are no versions in this table where the former GDR is the region of education).

Table 4.1 also offers some interesting insights into the relationship between "Abitur" shares and vocational wages. The even columns show that as for college wages, there is evidence of a negative relationship between "Abitur" shares and the wages received by those holding a vocational training degree exclusively. This may seem surprising at first glance, particularly in light of the results by Carneiro and Lee (2011) who find

[17]While a decline in average ability would call for improved matching mechanisms of children and young adults to the available educational tracks, a deterioration of the quality of upper secondary and university education could rather be counteracted, e.g., by raising the personnel and financial resources of these educational institutions.

[18]Figures C.6 and C.7 plot the evolution of real public expenditures for upper secondary schools (per graduate) by year and federal state, whereas figures C.8 and C.9 do the same for public expenditures for universities (per first-year student).

practically no effects of college shares on the wages received by those holding "only" a high school degree. Table 4.1, in contrast, provides tentative evidence of negative spillover effects of higher "Abitur" shares on sorting into vocational training. There are at least two alternative explanations for the existence of such effects: First, higher "Abitur" shares necessarily imply that relatively more young persons have the additional alternative to pursue a university degree instead of completing vocational training. As a result, the selection of applicants for vocational training may become somewhat less advantageous if the "marginal" college participants would have been among the top candidates for vocational training. Second, even if higher "Abitur" shares do not translate into correspondingly higher university shares, upper secondary schools may not prepare prospective vocational training participants as well for their careers as middle or lower secondary schools since the former are more academically oriented, whereas the latter two rather target vocationally inclined children and consequently convey somewhat more practical skills (Biewen and Tapalaga, 2017; Cockrill and Scott, 1997).

Similar to the case of college wages where only one specification excluding the GDR still reports a statistically significant estimate, the relationship between "Abitur" odds ratios and vocational wages is only significant when including the GDR as the region of education. However, the direction of the effect remains unchanged across the vast majority of specifications, and the coefficient estimates further remain similar in magnitude, ranging from about -0.6 to -0.3 when including the GDR and from roughly -0.4 to approximately zero when looking at West Germany only. Another interesting conclusion from table 4.1 concerns the way in which the coefficients in the equations for college and vocational wages relate to one another when the same set of controls is included: The detrimental impact of higher "Abitur" shares is always estimated to be considerably larger for college wages, implying that sorting into higher education has been more strongly affected than sorting into the medium-skilled group in terms of average ability.

Summing up, the results presented in table 4.1 lead to the following conclusions. First, they suggest that the expansion of upper secondary education has been accompanied by declining college wages – potentially caused by a less favourable selection of college graduates in terms of average ability –, with potential spillover effects on those holding a vocational training degree. Second, the detrimental effect of educational expansion is larger for college than for vocational wages. Completing the circle to what motivated the current research question in the first place, the next subsection explores the consequences of these observations for the college premium and its evolution over time.

4.6.2 Counterfactual college premia

As discussed above, it appears that educational expansion has affected both selection into tertiary education and into vocational training within educational groups, but that the effect on college wages was considerably larger. This implies that the college *premium* might have evolved differently in the absence of educational expansion. To illustrate this, it is interesting to consider a simple counterfactual thought experiment asking how the college premium would have evolved if "Abitur" shares were held fixed at the level of some base year b (see Carneiro and Lee, 2011). Remaining in the setting that measures average wages within cells, the experiment is based on the notion that the college premium in some year t can be obtained by aggregating across cells:

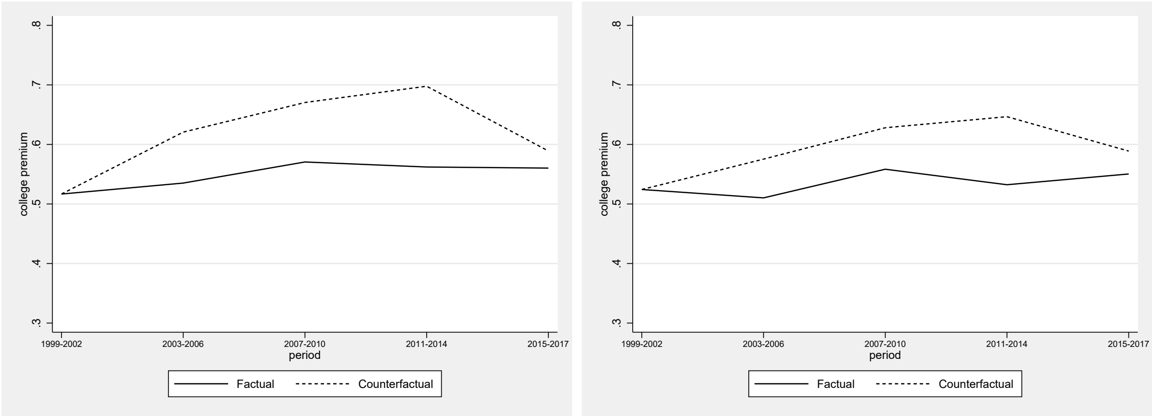
$$CP_t = \sum_{a=1}^{A}\sum_{r=1}^{R}\sum_{e=1}^{E}\omega_{atre}^{c} f_t^c(a,r,e) - \sum_{a=1}^{A}\sum_{r=1}^{R}\sum_{e=1}^{E}\omega_{atre}^{v} f_t^v(a,r,e) \tag{4.5}$$

where ω_{atre}^{c} and ω_{atre}^{v} denote average log college and vocational wages, respectively, and $f_t^c(a,r,e)$ and $f_t^v(a,r,e)$ denote the proportion of individuals in cell (a,r,e) for each skill type and year, which are used as weights for the aggregation. The college premium is then defined as the difference between aggregated average log college and vocational wages in the same year.

A "quality-adjusted" version of this college premium can be obtained by predicting average wages ω_{atre}^{k} based on a version of equation (4.4):

$$\omega_{atre}^{k} = \gamma_{atr}^{k} + \gamma_{ae}^{k} + \gamma_{te}^{k} + \gamma_{re}^{k} + \phi^k(P_{t-a,e}) + \lambda^k(P_{M,atre}^{k}, P_{M,atrr}^{k}) \quad ,$$

that, however, does not use the factual "Abitur" shares of each cohort $t-a$ from region e observed in year t. Instead, the shares are counterfactually held fixed at their level of a chosen base period b, meaning that the factual shares observed for someone from age group a in year t are replaced by those of age group a some years earlier (which were lower due to the observeed expansion in upper secondary schooling). This yields a college premium that would have been implied in year t if "Abitur" shares had stopped



(a) Including East Germany

(b) Excluding East Germany

Figure 4.6 – Aggregated college premium using factual and counterfactual average wages
(Source: NEPS-SC6-ADIAB, own calculations)

increasing in period b, i.e., if educational expansion had come to a halt at this point b:

$$CP_t^Q = \sum_{a=1}^{A}\sum_{r=1}^{R}\sum_{e=1}^{E}\omega_{atre}^{c,b} f_t^c(a,r,e) - \sum_{a=1}^{A}\sum_{r=1}^{R}\sum_{e=1}^{E}\omega_{atre}^{v,b} f_t^v(a,r,e) \quad , \tag{4.6}$$

with $\omega_{atre}^{k,b}$ denoting counterfactual wages for skill group k that would have resulted from the "Abitur" shares of year b, keeping everything else (skill prices etc.) at its factual level in year t. As also noted by Carneiro and Lee (2011), this counterfactual premium is not purged of selection, but rather the amount of selection into tertiary education and vocational training is held fixed at how it was in the chosen base period.

Since this thought experiment is based on the model in equation (4.4), it is only valid if the latter fits the data sufficiently well. As implied by the high R-squares in table 4.1 – amounting from about 90 to 93% for the vast majority of specifications –, figure C.5 in appendix C shows that the aggregated college premium based on predicted cell averages moves closely together with the factual premium observed in the data (obtained as the simple log difference between average college and vocational wages). In light of the differences between the specifications including and excluding the GDR and East Germany, respectively, the factual and counterfactual premia are computed and plotted against one another for both cases, where the counterfactuals are based on the estimates from the first and second columns (specifications (1/2a) and (1/2c), respectively). Both panels of figure 4.6 show that – according to the estimated model – the college premium would have been higher over the years if "Abitur" shares had remained at their level from around the millennium, where the discrepancy between factual and counterfactual premia is especially striking when including East Germany (see figure 4.6a). Even the somewhat more conservative estimates that are based exclusively on West Germany result in a counterfactual premium that is up to 10 percentage points higher than the factually observed one (see figure 4.6b). This is a sizeable effect whose magnitude should – however – be treated with caution given the limited number of observations.

4.7 Conclusion

The goal of this study was to explore whether the secular changes in educational attainment observed in Germany led to a changing composition particularly of university educated individuals in terms of their average ability. Using a combination of high-quality administrative wage data and rich survey data, the analysis exploits variation in educational shares between the regions where individuals went to secondary school in order to separate the wage effects of supply (which have already been studied in other contributions) from those of changes in quality (which have hitherto been neglected in the literature). The main findings can be summarised as follows.

First, the results suggest a negative relationship between college wages and the share of a cohort holding a university entry certificate. This is in line with the hypothesis that educational expansion is linked to a decline in the average ability of college graduates. While the estimated effects are economically meaningful, they are not always statistically significant, most likely due to the limited sample size. The latter also requires one to take the magnitude of the estimated effects with a grain of salt. Second, the analysis points to the existence of spillover effects on the medium-skilled group holding a vocational training degree, whose wages also tend to decline with higher shares of persons holding a university entry certificate. This can potentially be explained by a changing pool of applicants for vocational training as rising shares of young persons have the additional alternative to pursue a university degree. Third, selection effects appear to be larger for college than for vocational wages, meaning that the college *premium* would have been higher in absence of educational expansion. While these effects have reduced between-group inequality, it is likely that they explain part of the increased heterogeneity within groups found by other studies (see section 4.2). The findings of the current study are consistent with evidence from the US documenting a detrimental impact of educational expansion on the average ability of the highly skilled. Furthermore, they offer new insights on the inter-cohort differences in skill premia found in Germany by previous contributions that have hypothesised about, but not further explored the existence of quality effects.

The analysis also raises some questions that constitute interesting avenues for future research. First, the estimated quality effects can be understood as "gross effects" which are the sum of differences in "ex-ante quality" and changes in the quality of education due to educational expansion. Disentangling the two effects would be vital since the policy implications crucially differ depending on their individual importance. However, this would require detailed data on the financial *and* personnel means of educational institutions over time and between regions. Another issue left open by this study is to what extent educational expansion is related to dropout risk at universities. It seems likely that a decline in ability would – all else equal – lead to higher dropout rates (see also figure C.3), which may have long-lasting effects in terms of income and, hence, inequality (Zühlke et al., 2022). However, relating ability to dropout risk is challenging not least because grades may not be comparable over time or even between educational institutions (see Müller-Benedict and Gaens, 2020, for recent evidence on "grade inflation"). Isolating the effects of changing selection on dropout rates would substantially complement the results of the current study, which has thus far neglected this dimension.

Appendix C

C.1 Imputation of right-censored wages

As noted in section 4.5, wages in the SIAB are right-censored in the sense that employers only report them up to the social security contributions ceiling. Thus, denoting by w_{it} the true wage received by individual i in year t and letting c_t be the contributions ceiling, the observed wage is

$$w_i^{obs} = \begin{cases} w_{it} & \text{if} \quad w_{it} < c_t \\ c_t & \text{if} \quad w_{it} \geq c_t. \end{cases}$$

Wages amounting to at least 98 % of the contributions ceiling are therefore imputed based on Gartner (2005), who proposes to use a series of Tobit models estimated separately for each year and for East and West Germany, respectively. In the given application, the Tobit models include six categories of age, six categories of education, as well as their interactions. Since the expected values from the Tobit models suffer from overly high correlation with the covariates, the daily wages above the contributions ceiling are obtained by drawing from a truncated normal distribution, the lower limit of which is given by the respective contributions ceiling in year t.

C.2 Additional tables

Region	Federal states
1	Schleswig-Holstein (SH), Lower Saxony (NIE)
2	(West-)Berlin (BE), Hamburg (HH), Bremen (HB)
3	North Rhine-Westfalia (NRW), Hesse (HE)
4	Bavaria (BY), Baden-Wuerttemberg (BW), Saarland (SL), Rhineland-Palatinate (RPF)
5	Saxony (SN), Saxony-Anhalt (SA), Thuringia (TH), Brandenburg (BB), Mecklenburg-Western Pomerania (MV)
6 (only as region of education)	Former GDR

Table C.1 – Aggregation of federal states to regions of work and education

	Mean	Std. dev.	Min.	Max.
Conditioning on ≥ 5 observations per cell ($N = 576$)				
Average log college wages	5.247	0.331	4.263	6.071
Number of college observations per cell	31.368	45.961	5	249
Cohort “Abitur” share	0.173	0.074	0.038	0.355
Age group (a): 25-29	0.078	0.269	0	1
Age group (a): 30-34	0.144	0.351	0	1
Age group (a): 35-39	0.155	0.362	0	1
Age group (a): 40-44	0.148	0.350	0	1
Age group (a): 45-49	0.148	0.350	0	1
Age group (a): 50-54	0.142	0.349	0	1
Age group (a): 55-59	0.120	0.325	0	1
Age group (a): 60-64	0.066	0.248	0	1
Region of work (r): SH+NIE	0.170	0.376	0	1
Region of work (r): BE+HB+HH	0.248	0.432	0	1
Region of work (r): NRW+HE	0.238	0.426	0	1
Region of work (r): BY+BW+RPF+SL	0.247	0.431	0	1
Region of work (r): SN+SA+BB+MV+TH	0.097	0.297	0	1
Region of education (e): SH+NIE	0.212	0.409	0	1
Region of education (e): BE+HB+HH	0.099	0.299	0	1
Region of education (e): NRW+HE	0.243	0.429	0	1
Region of education (e): BY+BW+RPF+SL	0.165	0.371	0	1
Region of education (e): SN+SA+BB+MV+TH	0.063	0.242	0	1
Region of education (e): GDR	0.219	0.414	0	1
Years (t): 1999-2002	0.177	0.382	0	1
Years (t): 2003-2006	0.201	0.401	0	1
Years (t): 2007-2010	0.217	0.413	0	1
Years (t): 2011-2014	0.212	0.409	0	1
Years (t): 2015-2017	0.193	0.395	0	1
Migration rates	0.300	0.278	0.015	1
Staying rates	0.661	0.135	0.368	1
Conditioning on ≥ 10 observations per cell ($N = 380$)				
Average log college wages	5.257	0.299	4.263	5.880
Number of college observations per cell	44.937	52.258	10	249
Cohort “Abitur” share	0.168	0.667	0.061	0.323
Age group (a): 25-29	0.05	0.218	0	1
Age group (a): 30-34	0.145	0.352	0	1
Age group (a): 35-39	0.161	0.368	0	1
Age group (a): 40-44	0.150	0.358	0	1
Age group (a): 45-49	0.163	0.370	0	1
Age group (a): 50-54	0.158	0.365	0	1
Age group (a): 55-59	0.121	0.327	0	1
Age group (a): 60-64	0.053	0.224	0	1
Region of work (r): SH+NIE	0.158	0.365	0	1
Region of work (r): BE+HB+HH	0.203	0.402	0	1
Region of work (r): NRW+HE	0.268	0.443	0	1
Region of work (r): BY+BW+RPF+SL	0.274	0.446	0	1
Region of work (r): SN+SA+BB+MV+TH	0.097	0.297	0	1
Region of education (e): SH+NIE	0.234	0.424	0	1
Region of education (e): BE+HB+HH	0.063	0.244	0	1
Region of education (e): NRW+HE	0.242	0.429	0	1
Region of education (e): BY+BW+RPF+SL	0.184	0.388	0	1
Region of education (e): SN+SA+BB+MV+TH	0.055	0.229	0	1
Region of education (e): GDR	0.091	0.416	0	1
Years (t): 1999-2002	0.176	0.176	0	1
Years (t): 2003-2006	0.205	0.205	0	1
Years (t): 2007-2010	0.221	0.416	0	1
Years (t): 2011-2014	0.232	0.422	0	1
Years (t): 2015-2017	0.166	0.372	0	1
Migration rates	0.373	0.287	0.024	1
Staying rates	0.662	0.137	0.368	1

Table C.2 – Composition of cells entering regression of college wages

Note: The reported sample means of the dummy variables relating to age groups a, regions of work r, regions of education e, and periods t represent the fraction of cells belonging to each of these groups.
(Source: NEPS-SC6 ADIAB, own calculations)

	Mean	Std. dev.	Min.	Max.
Conditioning on $\geq$ 5 observations per cell ($N = 624$)				
Average log vocational wages	4.749	0.369	3.904	6.655
Number of vocational observations per cell	74.381	122.614	5	626
Cohort "Abitur" share	0.169	0.072	0.038	0.347
Age group (a): 25-29	0.088	0.283	0	1
Age group (a): 30-34	0.120	0.325	0	1
Age group (a): 35-39	0.133	0.340	0	1
Age group (a): 40-44	0.139	0.347	0	1
Age group (a): 45-49	0.144	0.352	0	1
Age group (a): 50-54	0.151	0.358	0	1
Age group (a): 55-59	0.146	0.353	0	1
Age group (a): 60-64	0.079	0.269	0	1
Region of work (r): SH+NIE	0.197	0.398	0	1
Region of work (r): BE+HB+HH	0.215	0.411	0	1
Region of work (r): NRW+HE	0.264	0.441	0	1
Region of work (r): BY+BW+RPF+SL	0.245	0.431	0	1
Region of work (r): SN+SA+BB+MV+TH	0.079	0.270	0	1
Region of education (e): SH+NIE	0.221	0.415	0	1
Region of education (e): BE+HB+HH	0.138	0.345	0	1
Region of education (e): NRW+HE	0.194	0.396	0	1
Region of education (e): BY+BW+RPF+SL	0.170	0.376	0	1
Region of education (e): SN+SA+BB+MV+TH	0.064	0.245	0	1
Region of education (e): GDR	0.213	0.410	0	1
Years (t): 1999-2002	0.191	0.393	0	1
Years (t): 2003-2006	0.208	0.406	0	1
Years (t): 2007-2010	0.215	0.411	0	1
Years (t): 2011-2014	0.206	0.405	0	1
Years (t): 2015-2017	0.179	0.384	0	1
Migration rates	0.299	0.273	0.004	0.971
Staying rates	0.782	0.100	0.588	0.971
Conditioning on $\geq$ 10 observations per cell ($N = 498$)				
Average log vocational wages	4.740	0.343	4.210	6.517
Number of vocational observations per cell	91.493	131.880	10	626
Cohort "Abitur" share	0.164	0.070	0.039	0.346
Age group (a): 25-29	0.086	0.281	0	1
Age group (a): 30-34	0.120	0.326	0	1
Age group (a): 35-39	0.135	0.342	0	1
Age group (a): 40-44	0.143	0.350	0	1
Age group (a): 45-49	0.153	0.360	0	1
Age group (a): 50-54	0.161	0.368	0	1
Age group (a): 55-59	0.129	0.335	0	1
Age group (a): 60-64	0.074	0.263	0	1
Region of work (r): SH+NIE	0.187	0.390	0	1
Region of work (r): BE+HB+HH	0.195	0.396	0	1
Region of work (r): NRW+HE	0.283	0.451	0	1
Region of work (r): BY+BW+RPF+SL	0.255	0.436	0	1
Region of work (r): SN+SA+BB+MV+TH	0.080	0.272	0	1
Region of education (e): SH+NIE	0.239	0.427	0	1
Region of education (e): BE+HB+HH	0.090	0.287	0	1
Region of education (e): NRW+HE	0.187	0.390	0	1
Region of education (e): BY+BW+RPF+SL	0.163	0.369	0	1
Region of education (e): SN+SA+BB+MV+TH	0.080	0.272	0	1
Region of education (e): GDR	0.241	0.428	0	1
Years (t): 1999-2002	0.189	0.392	0	1
Years (t): 2003-2006	0.205	0.404	0	1
Years (t): 2007-2010	0.217	0.413	0	1
Years (t): 2011-2014	0. 217	0.413	0	1
Years (t): 2015-2017	0.173	0.378	0	1
Migration rates	0.349	0.350	0.010	0.971
Staying rates	0.780	0.098	0.588	0.971

Table C.3 – Composition of cells entering regression of vocational wages

Note: The reported sample means of the dummy variables relating to age groups a, regions of work r, regions of education e, and periods t represent the fraction of cells belonging to each of these groups.
(Source: NEPS-SC6 ADIAB, own calculations)

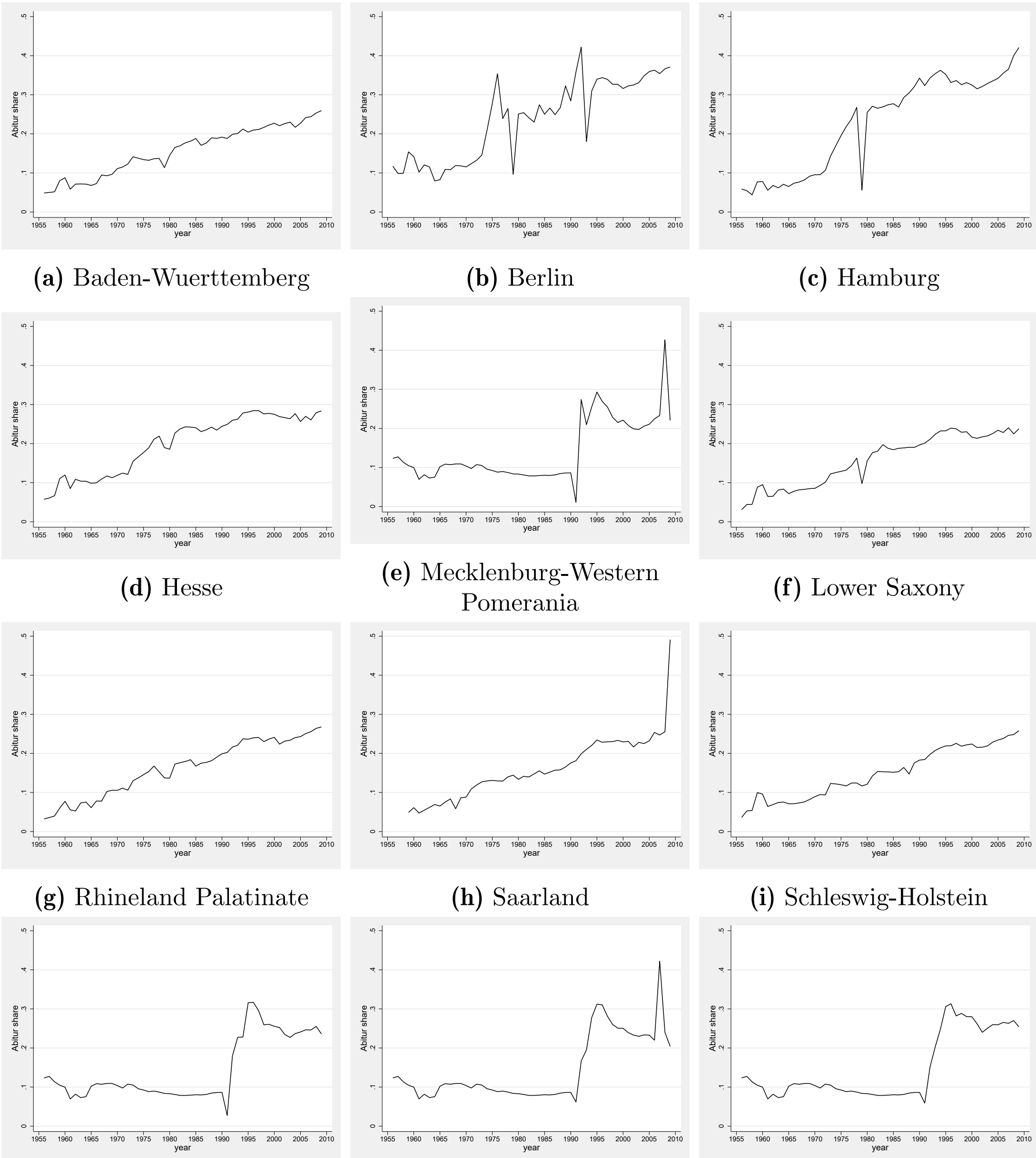
	(1) ln(wage^c)	(2) ln(wage^c)	(3) ln(wage^c)	(4) ln(wage^c)	(5) ln(wage^c)	(6) ln(wage^c)
(a) Including r = East Germany						
Odds ratio "Abitur"	-0.588	-0.772	-0.681	-0.740	-0.967*	-0.928
	[0.497]	[0.552]	[0.722]	[0.467]	[0.552]	[0.769]
School expenditures	0.005	-0.0002	0.008			
	[0.006]	[0.013]	[0.015]			
University expenditures				-0.004	-0.002	-0.002
				[0.003]	[0.003]	[0.003]
R^2	0.907	0.925	0.918	0.901	0.915	0.907
N	427	282	282	428	281	281
(b) Excluding r = East Germany						
Odds ratio "Abitur"	-0.586	-0.772	-0.681	-0.737	-0.967*	-0.928
	[0.457]	[0.554]	[0.711]	[0.457]	[0.544]	[0.758]
School expenditures	0.004	-0.0002	0.008			
	[0.006]	[0.012]	[0.015]			
University expenditures				-0.004	-0.002	-0.002
				[0.003]	[0.003]	[0.003]
R^2	0.897	0.915	0.907	0.930	0.915	0.907
N	401	273	273	403	273	273
Min. cell size	5	10	10	5	10	10
a-t-r fixed effects	YES	YES	YES	YES	YES	YES
e-a fixed effects	YES	YES	YES	YES	YES	YES
e-r fixed effects	YES	YES	NO	YES	YES	NO
e-t fixed effects	NO	NO	NO	NO	NO	NO
Migration rates	NO	NO	YES	NO	NO	YES

Table C.4 – Regressions including upper secondary school / university expenditures

(Source: NEPS-SC6-ADIAB, own calculations)

Note: Regressions of average college wages within (a, t, r, e)-cells (age, year, region of work, region of education) on "Abitur" odds ratio, public expenditures for upper secondary schools per graduate student (columns (1)-(3)) and for universities per first-year student (columns (4)-(6)) given for cells defined by $(t - a, e)$ (year–age, region of education). Observations (i.e. cells) are weighted by the inverse sampling variance of log college wages within cells. Heteroskedasticity-robust standard errors in parentheses.
*: significant at 10 % level; **: significant at 5 % level; ***: significant at 1 % level.

C.3 Additional figures



(a) Baden-Wuerttemberg **(b)** Berlin **(c)** Hamburg

(d) Hesse **(e)** Mecklenburg-Western Pomerania **(f)** Lower Saxony

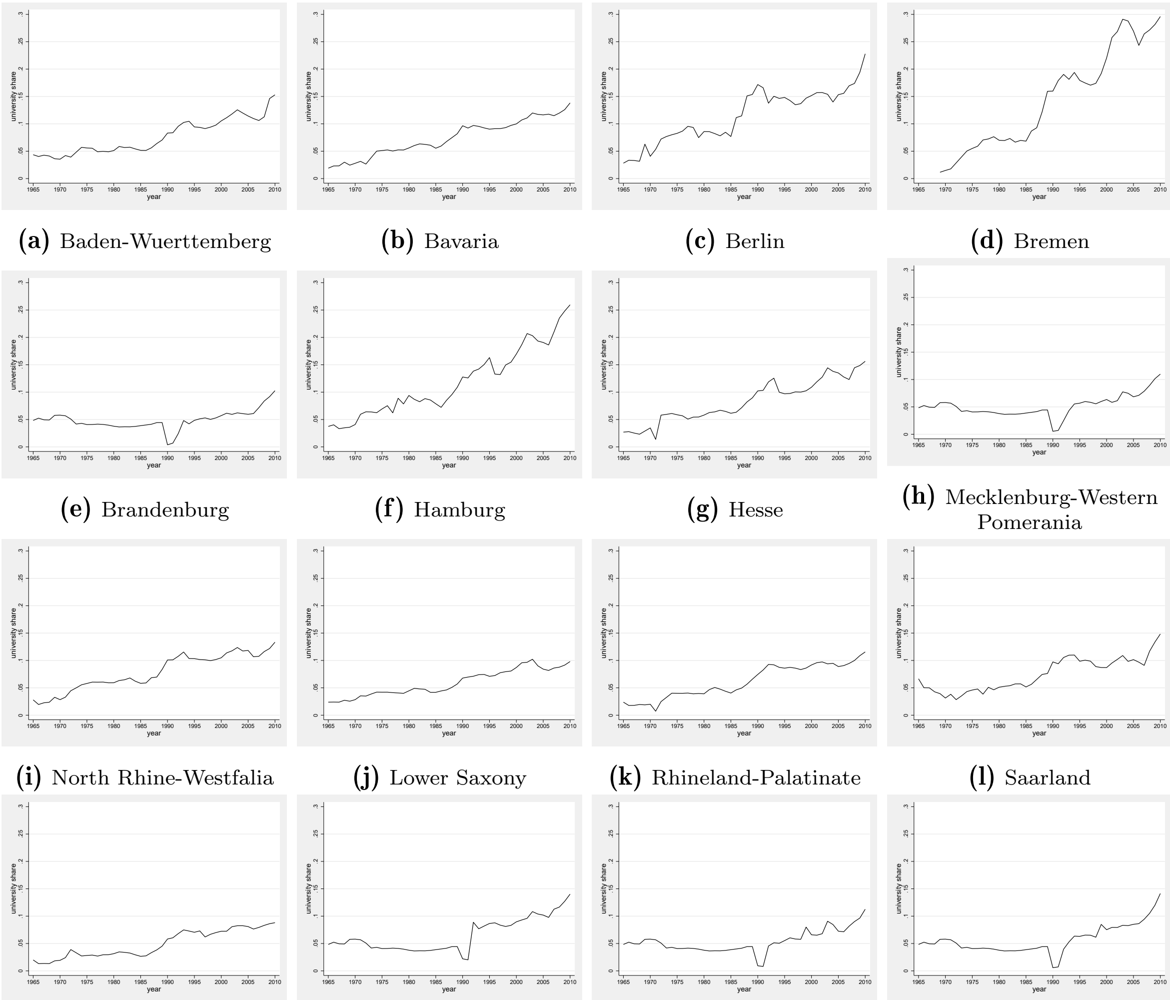
(g) Rhineland Palatinate **(h)** Saarland **(i)** Schleswig-Holstein

(j) Saxony **(k)** Saxony-Anhalt **(l)** Thuringia

Figure C.1 – "Abitur" share by federal state

(Source: Statistical Yearbooks, Federal Statistical Office, own calculations)

Note: The "Abitur" share is computed as the number of persons graduating with a university entry certificate divided by the size of the relevant age group (18-21 years of age) in the respective federal state.



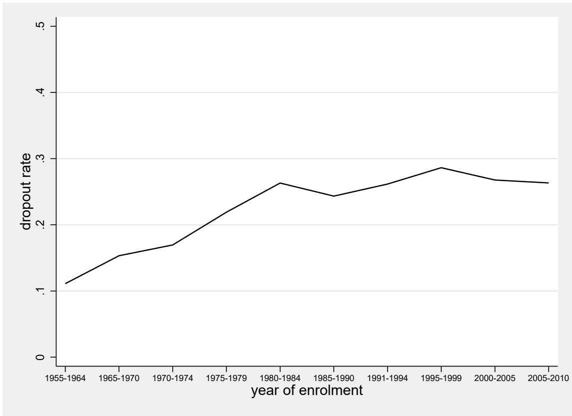
(a) Baden-Wuerttemberg **(b)** Bavaria **(c)** Berlin **(d)** Bremen

(e) Brandenburg **(f)** Hamburg **(g)** Hesse **(h)** Mecklenburg-Western Pomerania

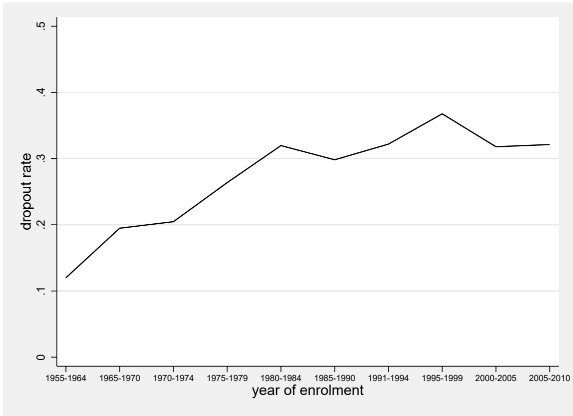
(i) North Rhine-Westfalia **(j)** Lower Saxony **(k)** Rhineland-Palatinate **(l)** Saarland

(m) Schleswig-Holstein **(n)** Saxony **(o)** Saxony-Anhalt **(p)** Thuringia

Figure C.2 – First-year student ratio by federal state
(Source: Statistical Yearbooks, Federal Statistical Office, own calculations)
Note: The first-year student ratio is computed as the number of first-year students divided by the size of the relevant age group (18-21 years of age) in the respective federal state.

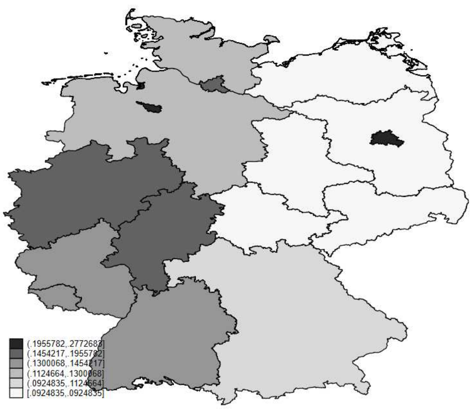


(a) University & “Fachhochschule”

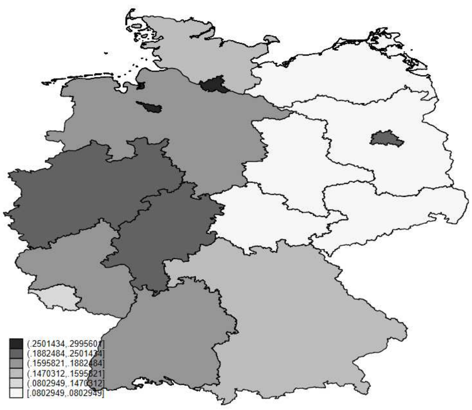


(b) University only

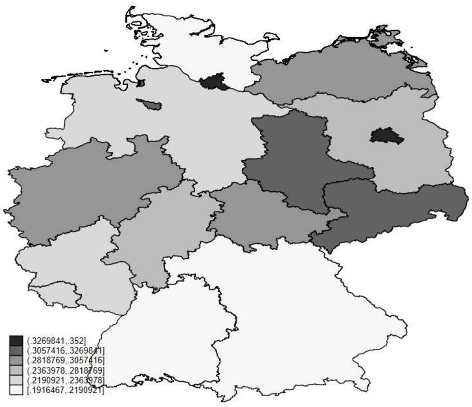
Figure C.3 – University and FH dropout rates by year of enrolment
(Source: NEPS-SC6 v11, own calculations)



(a) 1975

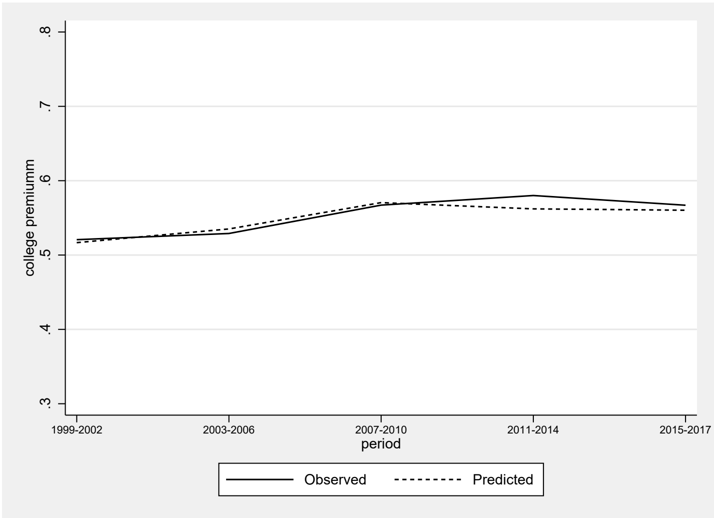


(b) 1985

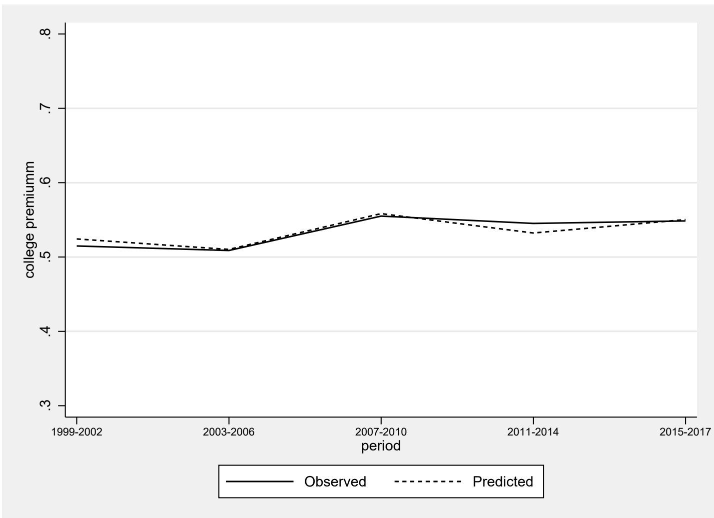


(c) 1995

Figure C.4 – Shares of young persons graduating with “Abitur” in the 16 federal states
(Source: Statistical Yearbooks, Federal Statistical Office)

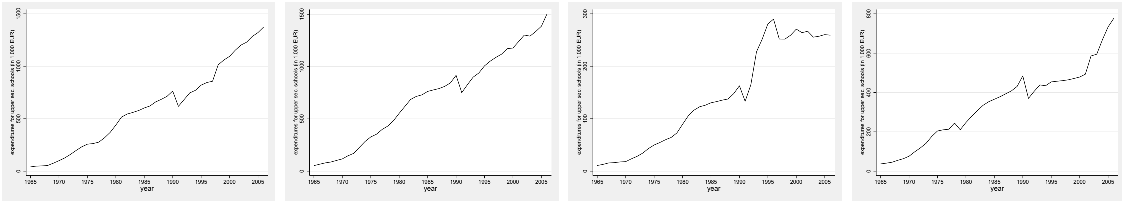


(a) Including GDR & East Germany

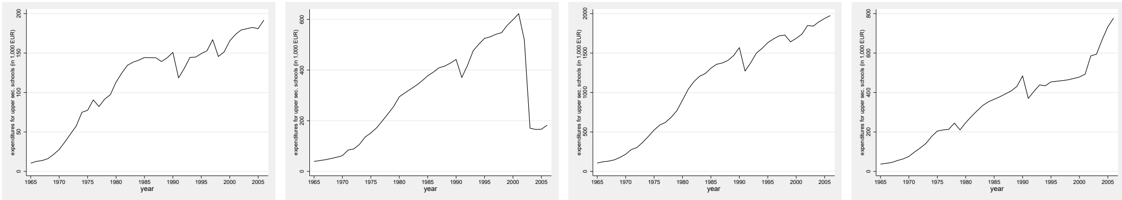


(b) Excluding GDR & East Germany

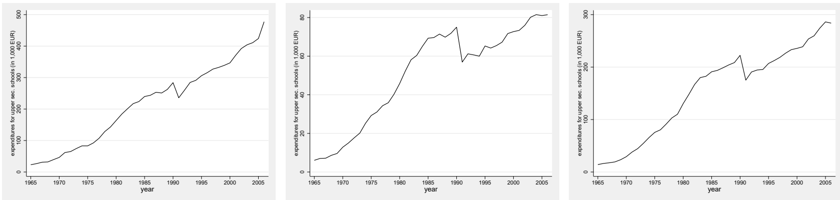
Figure C.5 – Aggregated college premium using observed and predicted average wages
(Source: NEPS-SC6-ADIAB, own calculations)
Note: The predicted premia (dashed line) are based on specifications (1a) and (2a) and (1c) and (2c) in table 4.1, respectively.



(a) Baden-Wuerttemberg **(b)** Bavaria **(c)** Berlin **(d)** Bremen

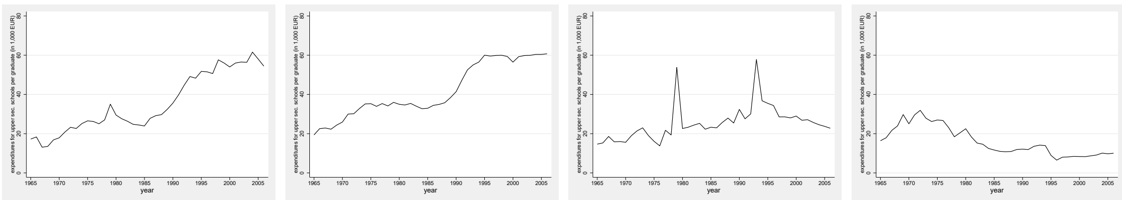


(e) Hamburg **(f)** Hesse **(g)** North Rhine-Westfalia **(h)** Lower Saxony

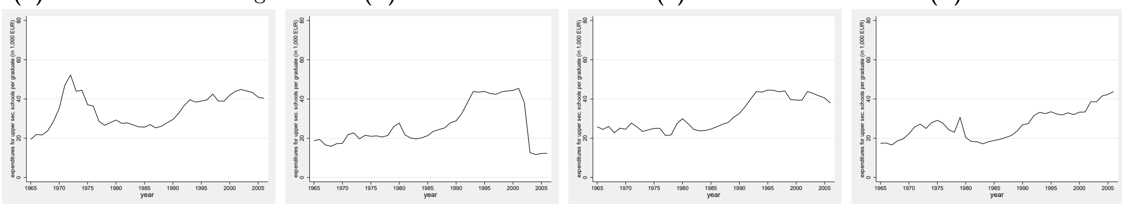


(i) Rhineland-Palatinate **(j)** Saarland **(k)** Schleswig-Holstein

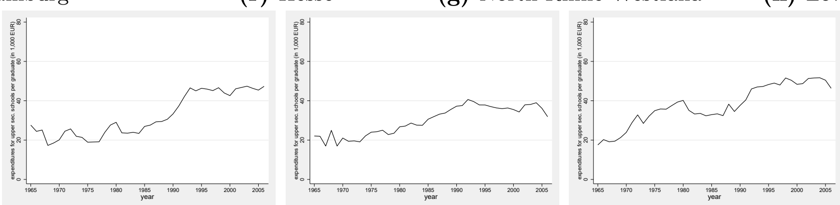
Figure C.6 – Real expenditures for upper secondary schools by federal state
(Source: Statistical Yearbooks, Federal Statistical Office)
Note: Panel (c) shows data for West-Berlin only prior to 1992.



(a) Baden-Wuerttemberg **(b)** Bavaria **(c)** Berlin **(d)** Bremen

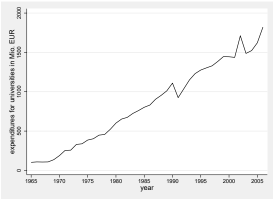
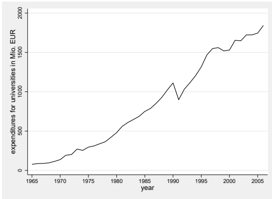
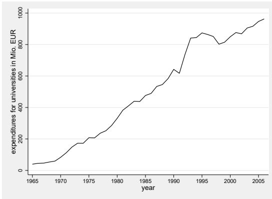
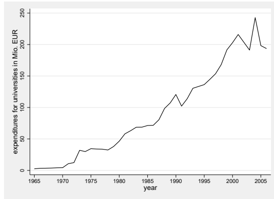


(e) Hamburg **(f)** Hesse **(g)** North Rhine-Westfalia **(h)** Lower Saxony

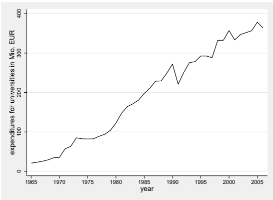
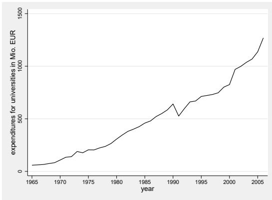
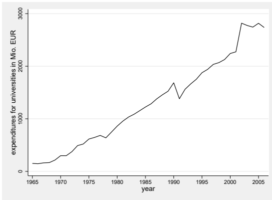
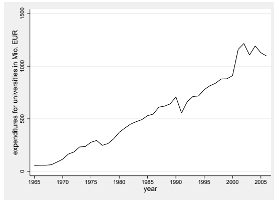


(i) Rhineland-Palatinate **(j)** Saarland **(k)** Schleswig-Holstein

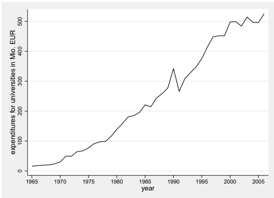
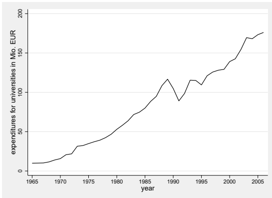
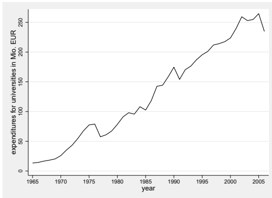
Figure C.7 – Real expenditures per upper secondary school graduate by federal state
(Source: Statistical Yearbooks, Federal Statistical Office, own calculations)
Note: Panel (c) shows data for West-Berlin exclusively prior to 1992.

(a) Baden-Wuerttemberg **(b)** Bavaria **(c)** Berlin **(d)** Bremen

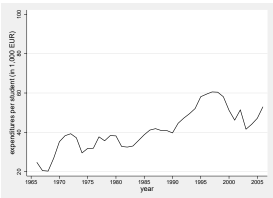
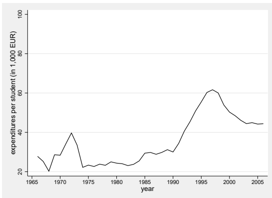
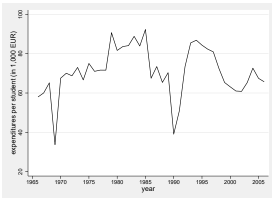
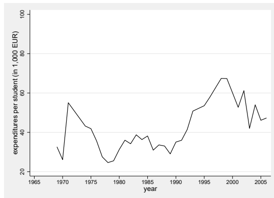
   

(e) Hamburg **(f)** Hesse **(g)** North Rhine-Westfalia **(h)** Lower Saxony

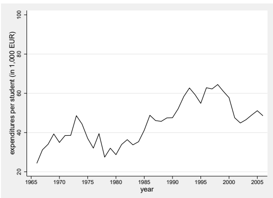
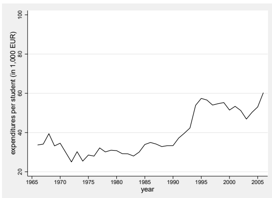
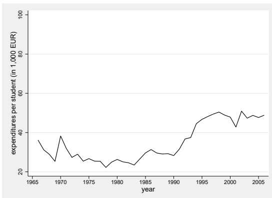
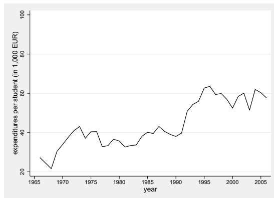
  

(i) Rhineland-Palatinate **(j)** Saarland **(k)** Schleswig-Holstein

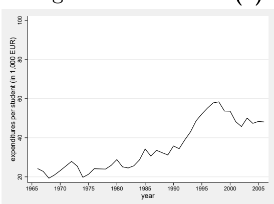
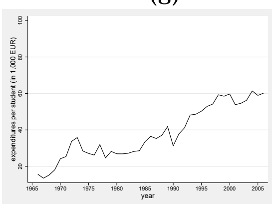
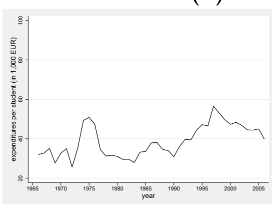
Figure C.8 – Real expenditures for universities by federal state
(Source: Statistical Yearbooks, Federal Statistical Office)
Note: Panel (c) shows data for West-Berlin exclusively prior to 1992.

(a) Baden-Wuerttemberg **(b)** Bavaria **(c)** Berlin **(d)** Bremen

(e) Hamburg **(f)** Hesse **(g)** North Rhine-Westfalia **(h)** Lower Saxony

(i) Rhineland-Palatinate **(j)** Saarland **(k)** Schleswig-Holstein

Figure C.9 – Real expenditures for universities per first-year student by federal state
(Source: Statistical Yearbooks, Federal Statistical Office)
Note: Panel (c) shows data for West-Berlin only prior to 1992.

CHAPTER 5

Dissertation conclusion

Since the millennium, Germany's economy – and the German labour market in particular – has undergone major changes: Once denoted as the "sick man of Europe", Germany saw an unparalleled economic boom starting in the mid 2000s that was only interrupted by the COVID-19 pandemic in 2020. Likewise, the first two decades of the 21st century were characterised by a trend reversal in income inequality, which had been climbing steadily since the reunification, but finally stagnated (or even slightly declined depending on the precise income measure) after 2010. Motivated by these observations, this dissertation aims to contribute to a better understanding of the link between documented trends in inequality and secular trends on the German labour market.

Setting the scene for the subsequent studies, chapter 2 of this thesis presents a rigorous decomposition analysis breaking down the observed distributional changes in net equivalised household incomes between 2005/06 and 2015/16 into the contributions of eight factors overall. We thereby focus in particular on the role played by the striking expansion in employment over this period, which should (all else equal) have primarily benefitted individuals at the bottom of the distribution, but has not translated into a corresponding decline in the commonly used inequality indices. The results of our analysis reconcile these seemingly contradictive observations, showing that the employment boom did have a significant equalising effect that was, however, masked by the countervailing impact of other factors, most notably immigration (resulting in an inflow of tentatively low-income individuals) and improved household characteristics such as educational attainment. While these have led to gains in disposable income across the entire distribution, their positive effect is highly non-uniform and becomes more pronounced towards the top percentiles, thus contributing to an increase in inequality. Our analysis further reveals that the *gross* effect of the employment boom would have led to considerably larger gains at the bottom of the distribution, but was significantly dampened by Germany's tax and transfer system. If reducing inequality is a socially desired target, this may seem somewhat sobering. On the bright side, however, our analysis also demonstrates the effectiveness of the German redistributive system, which can conversely be expected to dampen inequality-*increasing* effects during recessions. In light of the economic challenges currently faced by Germany – including rising unemployment after more than a decade of a booming labour market –, this is a highly relevant and encouraging result.

Although chapter 2 is not exclusively concerned with labour income, the main results of the analysis are tightly connected to two of the most salient trends on the German labour market: First, the rising employment levels driven in particular by increasing participation rates of women; and second, the ongoing educational expansion (often denoted as an "academisation") of the German workforce. The remaining two chapters of this dissertation investigate the evolution of two wage differentials related to these trends: the gender wage gap (see chapter 3) and the college premium (see chapter 4).

Against the background of the substantial changes in labour market participation – potentially altering the composition of the workforce –, chapter 3 reassesses the German gender wage gap and its evolution from the recession (2000-2005) to the peak of the employment boom (2012-2017), focussing in particular on sample selection. Selection on unobservables is a widely known complication in the context of the gender wage gap, causing biased measures of the gap if men and women are unequally affected by it. Although the existing literature proposes a vast variety of alternative selectivity-corrected measures, many of these ignore the potential for male selection bias and – most importantly – they impose homogeneous selection on unobservables throughout the wage distribution. Chapter 3, in contrast, explicitly allows for heterogeneous selection patterns by exploiting the novel distribution regression framework established by Chernozhukov et al. (2025). Our results indeed reveal substantial heterogeneities: For full-time men, we find positive selection at the bottom – possibly related to Germany's generous social safety net –, whereas it is negative in the rest of the distribution. Full-time women, on the other hand, appear to be generally negatively selected, and increasingly so towards the top. This may be driven by assortative matching, particularly in combination with the German tax and transfer system in which the secondary earner typically faces high marginal tax rates. The patterns of male and female selectivity explain a large part of the wage differences between men and women, although they have somewhat converged between 2000 and 2017: Whereas men have become less favourably selected over the course of the employment boom, women's negative selectivity has become slightly weaker, likely due to changing social norms or improvements in public childcare. For part-time employment, we find generally negative selectivity for men and a highly non-uniform pattern for women, turning from negative in the lower half to highly positive in the upper part. Analogous to the full-time case, these patterns somewhat converge over time, contributing once more to a decline in the respective gender wage gap. All in all, our analysis presents comprehensive evidence of the heterogeneity of both female and male selection into employment, challenging the results obtained by previous studies based on more restrictive analytical settings.

Chapter 4 finally turns the attention to the ongoing educational expansion (or "academisation") of the German workforce, which constitutes perhaps the most striking trend on the labour market apart from the record employment levels. The continuously rising tertiary education shares over the last decades have been accompanied by significant differences in skill premia across cohorts, which has inspired discussions as to whether these differences can be explained by a declining (academic) ability of the highly educated. Showing this empirically is challenging since changes in educational shares translate into changes in relative supplies, making it hard to disentangle quality effects from supply effects. Exploiting the federal organisation of the German education system, chapter 4 resolves the issue by exploiting *regional* variation in educational shares, comparing the wages of highly educated persons who received their university entry certificate ("Abitur") in different German states, but who otherwise face the same labour market conditions. The results of this analysis tend to imply decreasing college wages with higher shares of persons formally qualified for tertiary education, suggesting that educational expansion is linked to a decline in the average ability of the highly educated. The analysis further points to spillover effects of educational expansion on the medium-skilled group, possibly explained by a smaller pool of applicants as more and more individuals face the additional alternative to go to college. However, the adverse effects linked to the quality of a cohort are not as large as for the highly educated group, such that – as illustrated by a simple counterfactual exercise – the college premium would have been higher in absence of the observed educational expansion.

In conclusion, the contribution of this dissertation can be summarised as follows. From a methodical perspective, chapters 2 and 3 underscore the importance of taking a distributional standpoint in the analysis of income inequality. Notwithstanding the usefulness of indices such as the Gini in providing a first picture of general inequality trends, the informational value of these measures is often limited in the sense that it remains obscure which parts of the distribution are responsible for the observed changes. For instance, chapter 2 reveals that the apparent stagnation in disposable income inequality (as measured by the Gini) hides important information such as the general, but non-uniform relative growth of incomes across the entire distribution. Apart from this, our analysis demonstrates that different developments affect different parts of the distribution unequally, which may bear important implications for policymakers. Chapter 3 further shows that potential biases caused by selection on unobservables may also be heterogeneous across the distribution, a possibility that has been widely ignored thus far. Thanks to the aforementioned advances in the econometric literature on selection (see Chernozhukov et al., 2025), the opportunity to explore such heterogeneities has recently opened up to researchers. This dissertation is among the first contributions to provide corresponding evidence. On a more general note, chapters 3 and 4 both point to the importance of unobservables in measuring wage differentials such as the gender wage gap or the college premium and their evolution over time. Another conclusion from this dissertation is therefore that any analysis of observed trends in between-group inequalities should be aware of potential changes in the composition of unobservables among the groups considered, and apply suitable correction techniques if necessary.

Beside these methodical considerations, the individual chapters this thesis is based upon are all concerned with topics that are as present as ever not only in academia, but also in political and public discourse. As an example, the incessantly rising teritiary education shares addressed in chapter 4 are controversially discussed in light of two simultaneous developments: the increasing skills shortage in vocational professions – intensified by the pending retirement of the baby boomer generation (Deschermeier and Schäfer, 2024) –, and the revolutionary advances in Artificial Intelligence (AI) affecting predominantly routine cognitive tasks often executed by college graduates (Eloundou et al., 2024; Green, 2024). Against this background, future cohorts may make different career choices than the preceeding ones,[1] perhaps leading to a turnaround in the observed educational trends. This question as well as the extent to which AI will replace or complement skills in the future – with corresponding effects on skill premia – will certainly continue to be a concern both inside and outside of academia in the years to come. Likewise, the issue of inequality or – more broadly speaking – social justice remains among the most frequently recurring themes in public debates, no matter whether in the context of the gender wage gap, the German pension system, or in debates surrounding unemployment benefits (the so-called "Bürgergeld" in particular). Discussions of this nature often tend to be emotionally charged since social justice is – as already mentioned in the introduction of this thesis – inherently connected to highly subjective normative values. For this reason, a constructive public and political discourse on inequality would not be conceivable without a profound empirical basis. A final (and crucial) goal of this dissertation is therefore to contribute to further building this basis.

[1]Goller et al. (2025) present some first evidence on the changing occupational search behaviour of high school graduates in Switzerland, which is characterised by a similar dual vocational training system as Germany.

Bibliography

Acemoglu, Daron (2002). Technical change, inequality, and the labor market. *Journal of Economic Literature* 40 (1), pages 7–72.

Acemoglu, Daron and David Autor (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of Labor Economics*, Volume 4, pages 1043–1171. Elsevier.

Akyol, Metin, Michael Neugart and Stefan Pichler (2013). Were the Hartz reforms responsible for the improved performance of the German labour market? *Economic Affairs* 33 (1), pages 34–47.

Albers, Thilo N. H., Charlotte Bartels and Moritz Schularick (2025). Wealth and its distribution in Germany, 1895-2021. *Journal of the European Economic Association*, forthcoming.

Albrecht, James, Aico Van Vuuren and Susan Vroman (2009). Counterfactual distributions with sample selection adjustments: Econometric theory and an application to the Netherlands. *Labour Economics* 16 (4), pages 383–396.

Alesina, Alberto and Eliana La Ferrara (2000). Participation in heterogeneous communities. *The Quarterly Journal of Economics* 115 (3), pages 847–904.

Allmendinger, Jutta (1989). Educational systems and labor market outcomes. *European Sociological Review* 5 (3), pages 231–250.

Ammermüller, Andreas and Andrea Maria Weber (2005). Educational attainment and returns to education in Germany. *ZEW-Centre for European Economic Research Discussion Paper* 05-017.

Antonczyk, Dirk, Thomas DeLeire and Bernd Fitzenberger (2018). Polarization and rising wage inequality: Comparing the US and Germany. *Econometrics* 6 (2), pages 1–33.

Antoni, Manfred, Andreas Ganzer and Philipp vom Berge (2019). Sample of Integrated Labour Market Biographies Regional File (SIAB-R) 1975 - 2017. *FDZ-Datenreport* 04/2019.

Arellano, Manuel and Stéphane Bonhomme (2017). Quantile selection models with an application to understanding changes in wage inequality. *Econometrica* 85 (1), pages 1–28.

Armour, Philip, Richard V. Burkhauser and Jeff Larrimore (2013). Deconstructing income and income inequality measures: A crosswalk from market income to comprehensive income. *American Economic Review* 103 (3), pages 173–177.

Atkinson, Anthony B. (2015). *Inequality: What can be done?* Harvard University Press.

Autor, David H., Lawrence F. Katz and Melissa S. Kearney (2008). Trends in U.S. wage inequality: Revising the revisionists. *The Review of Economics and Statistics* 90 (2), pages 300–323.

Bach, Stefan, Peter Haan and Richard Ochmann (2013). Taxation of married couples in Germany and the UK: One-earner couples make the difference. *International Journal of Microsimulation* 6 (3), pages 2–20.

Bachbauer, Nadine, Clara Wolf and Daniel Fuß (2021). Education and employment trajectories in NEPS-ADIAB: The survey data of the National Educational Panel Study Linked to Administrative Data of the Institute for Employment Research. *European Sociological Review* 38 (4), pages 663–676.

Baier, Tina and Marcel Helbig (2014). Much ado about EUR 500: Do tuition fees keep German students from entering university? Evidence from a natural experiment using DiD matching methods. *Educational Research and Evaluation* 20 (2), pages 98–121.

Bardasi, Elena and Janet C. Gornick (2008). Working for less? Women's part-time wage penalties across countries. *Feminist Economics* 14 (1), pages 37–72.

Bartels, Charlotte and Katharina Jenderny (2015). The role of capital income for top income shares in Germany. *WTID Working Paper* 1-2015.

Batruch, Anatolia, Sara Geven, Emma Kessenich and Herman G. van de Werfhorst (2023). Are tracking recommendations biased? A review of teachers' role in the creation of inequalities in tracking decisions. *Teaching and Teacher Education* 123, page 103985.

Bauer, Thomas, Holger Bonin, Lorenz Goette and Uwe Sunde (2007). Real and nominal wage rigidities and the rate of inflation: Evidence from West German micro data. *Economic Journal* 117 (524), pages 508–529.

Baumgarten, Daniel, Gabriel Felbermayr and Sybille Lehwald (2020). Dissecting between-plant and within-plant wage dispersion: Evidence from Germany. *Industrial Relations: A Journal of Economy and Society* 59 (1), pages 85–122.

Beblo, Miriam, Denis Beninger, Anja Heinze and François Laisney (2003). Measuring selectivity-corrected gender wage gaps in the EU. *ZEW Discussion Paper* (03-74).

Bick, Alexander (2015). The quantitative role of child care for female labor force participation and fertility. *Journal of the European Economic Association* 14 (3), pages 639–668.

Bick, Alexander and Nicola Fuchs-Schündeln (2017). Quantifying the disincentive effects of joint taxation on married women's labor supply. *American Economic Review, P&P* 107 (5), pages 100–104.

Biewen, Martin (2002). Bootstrap inference for inequality, mobility and poverty measurement. *Journal of Econometrics* 108, pages 317 –342.

Biewen, Martin, Bernd Fitzenberger and Matthias Seckler (2020). Counterfactual quantile decompositions with selection correction taking into account Huber/Melly (2015): An application to the German gender wage gap. *Labour Economics* 67, page 101927.

Biewen, Martin and Andos Juhasz (2012). Understanding rising income inequality in Germany, 1999/2000–2005/2006. *Review of Income and Wealth* 58 (4), pages 622–647.

Biewen, Martin and Matthias Seckler (2019). Unions, internationalization, tasks, firms, and worker characteristics: A detailed decomposition analysis of rising wage inequality in Germany. *Journal of Economic Inequality* 17, pages 461–498.

Biewen, Martin and Miriam Sturm (2022). Why a labour market boom does not necessarily bring down inequality: Putting together Germany's inequality puzzle. *Fiscal Studies* 43 (2), pages 121–149.

Biewen, Martin and Madalina Tapalaga (2017). Life-cycle educational choices in a system with early tracking and 'second chance'options. *Economics of Education Review* 56, pages 80–94.

Biewen, Martin and Madalina Thiele (2020). Early tracking, academic vs. vocational training, and the value of 'second-chance' options. *Labour Economics* 66, page 101900.

Biewen, Martin, Martin Ungerer and Max Löffler (2019). Why did income inequality in Germany not increase further after 2005? *German Economic Review* 20 (4), pages 471–504.

Blanden, Jo (2013). Cross-country rankings in intergenerational mobility: a comparison of approaches from economics and sociology. *Journal of Economic Surveys* 27 (1), pages 38–73.

Blau, Francine D., Lawrence M. Kahn, Nikolai Boboshko and Matthew Comey (2024). The impact of selection into the labor force on the gender wage gap. *Journal of Labor Economics* 42 (4), pages 1093–1133.

Blinder, Alan S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human Resources* 8 (4), pages 436–455.

Blundell, Richard, Amanda Gosling, Hide Ichimura and Costas Meghir (2007). Changes in the distribution of male and female wages accounting for the employment composition using bounds. *Econometrica* 75 (2), pages 323 –363.

Blundell, Richard, Robert Joyce, Agnes Norris Keiller and James P. Ziliak (2018). Income inequality and the labour market in Britain and the US. *Journal of Public Economics* 162, pages 48 –62.

Bol, Thijs and Herman G. Van de Werfhorst (2013). Educational systems and the trade-off between labor market allocation and equality of educational opportunity. *Comparative Education Review* 57 (2), pages 285–308.

Bonin, Holger, Annabelle Krause-Pilatus and Ulf Rinne (2020). Selbständige Erwerbstätigkeit in Deutschland (Aktualisierung 2020). Forschungsbericht 545. Bundesministerium für Arbeit und Soziales, Berlin.

Boockmann, Bernhard and Viktor Steiner (2006). Cohort effects and the returns to education in West Germany. *Applied Economics* 38 (10), pages 1135–1152.

Bossler, Mario and Thorsten Schank (2023). Wage inequality in Germany after the minimum wage introduction. *Journal of Labor Economics* 41 (3), pages 813–857.

Bradley, Jake and Alice Kügler (2019). Labor market reforms: An evaluation of the Hartz policies in Germany. *European Economic Review* 113, pages 108 –135.

Bruckmeier, Kerstin and Berthold U. Wigger (2014). The effects of tuition fees on transition from high school to university in Germany. *Economics of Education Review* 41, pages 14–23.

Brüll, Eduard and Christina Gathmann (2020). Evolution of the East German wage structure. *ZEW-Centre for European Economic Research Discussion Paper* (20-081).

Brunello, Giorgio and Daniele Checchi (2007). Does school tracking affect equality of opportunity? New international evidence. *Economic Policy* 22 (52), pages 782–861.

Buchinsky, Moshe (1998). The dynamics of changes in the female wage distribution in the USA: a quantile regression approach. *Journal of Applied Econometrics* 13 (1), pages 1–30.

Buchinsky, Moshe (2001). Quantile regression with sample selection: Estimating women's return to education in the US. *Economic Applications of Quantile Regression*, pages 87–113.

Bundesamt für Migration und Flüchtlinge (2017). Migrationsbericht der Bundesregierung 2016/17.

Burda, Michael (2016). The German labor market miracle, 2003 -2015: An assessment. *SFB Discussion Papers* 2006-005.

Burda, Michael and Stefanie Seele (2020). Reevaluating the German labor market miracle. *German Economic Review* 21 (2), pages 139–179.

Calvo, Paula, Ilse Lindenlaub and Ana Reynoso (2024). Marriage market and labour market sorting. *Review of Economic Studies* 91 (6), pages 3316–3361.

Card, David, Jörg Heining and Patrick Kline (2013). Workplace heterogeneity and the rise of West German wage inequality. *The Quarterly Journal of Economics* 128 (3), pages 967–1015.

Card, David and Thomas Lemieux (2001). Can falling supply explain the rising return to college for younger men? A cohort-based analysis. *The Quarterly Journal of Economics* 116 (2), pages 705–746.

Carneiro, Pedro and Sokbae Lee (2011). Trends in quality-adjusted skill premia in the United States, 1960-2000. *American Economic Review* 101 (6), pages 2309–49.

Carrillo-Tudela, Carlos, Andrey Launov and Jean-Marc Robin (2021). The fall in German unemployment: A flow analysis. *European Economic Review* 132, page 103658.

Chen, Songnian, Nianqing Liu, Hanghui Zhang and Yahong Zhou (2024). Estimation of wage inequality in the UK by quantile regression with censored selection. *Journal of Econometrics*, page 105733.

Chernozhukov, Victor, Denis Chetverikov and Kengo Kato (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *The Annals of Statistics* 41 (6), pages 2786–2819.

Chernozhukov, Victor, Ivan Fernandez-Val and Siyi Luo (2025). Distribution regression with sample selection and UK wage decomposition. *Journal of Political Economy,* forthcoming.

Chernozhukov, Victor, Ivan Fernandez-Val, Blaise Melly and Kaspar Wüthrich (2020). Generic inference on quantile and quantile effect functions for discrete outcomes. *Journal of the American Statistical Society* 115 (529), pages 123–127.

Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron and David Cutler (2016). The association between income and life expectancy in the United States, 2001-2014. *Journal of the American Medical Association* 315 (16), pages 1750–1766.

Chzhen, Yekaterina and Karen Mumford (2011). Gender gaps across the earnings distribution for full-time employees in Britain: Allowing for sample selection. *Labour Economics* 18 (6), pages 837–844.

Cockrill, Antje and Peter Scott (1997). Vocational education and training in Germany: Trends and issues. *Journal of Vocational Education and Training* 49 (3), pages 337–350.

Corak, Miles (2013). Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives* 27 (3), pages 79–102.

Cowell, Frank A. (2000). Measurement of inequality. In *Handbook of Income Distribution*, Volume 1, pages 87–166. Elsevier.

Cunha, Flavio, Fatih Karahan and Ilton Soares (2011). Returns to skills and the college premium. *Journal of Money, Credit and Banking* 43 (1), pages 39–86.

Dahlberg, Matz and Magnus Gustavsson (2008). Inequality and crime: separating the effects of permanent and transitory income. *Oxford Bulletin of Economics and Statistics* 70 (2), pages 129–153.

Daly, Mary C. and Robert G. Valletta (2006). Inequality and poverty in United States: the effects of rising dispersion of men's earnings and changing family behaviour. *Economica* 73 (289), pages 75–98.

Dauth, Wolfgang, Sebastian Findeisen and Jens Suedekum (2017). Trade and manufacturing jobs in Germany. *American Economic Review* 107 (5), pages 337–342.

Dauth, Wolfgang, Sebastian Findeisen and Jens Suedekum (2021). Adjusting to globalization in Germany. *Journal of Labor Economics* 39 (1), pages 1–30.

Deschermeier, Philipp and Holger Schäfer (2024). Die Babyboomer gehen in Rente. *IW-Kurzbericht* (78).

Dillon, Eleanor W. and Gregory Veramendi (2018). The role of sorting and skill prices in the evolution of the college premium. *Working Paper*.

DiNardo, John, Nicole Fortin and Thomas Lemieux (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica* 64 (5), pages 1011–1044.

Doepke, Matthias and Ruben Gaetani (2024). Why didn't the college premium rise everywhere? Employment protection and on-the-job investment in skills. *American Economic Journal: Macroeconomics* 16 (3), pages 268–309.

Dolado, Juan J., Cecilia García-Peñalosa and Linas Tarasonis (2020). The changing nature of gender selection into employment over the great recession. *Economic Policy* 35 (104), pages 635–677.

Drechsel-Grau, Moritz, Andreas Peichl and Kai D. Schmid (2015). Einkommensverteilung und gesamtwirtschaftliche Entwicklung in Deutschland. *Wirtschaftsdienst* 95, pages 684–688.

Drechsel-Grau, Moritz, Andreas Peichl, Kai D. Schmid, Johannes F. Schmieder, Hannes Walz and Stefanie Wolter (2022). Inequality and income dynamics in Germany. *Quantitative Economics* 13 (4), pages 1593–1635.

Dustmann, Christian, Bernd Fitzenberger, Uta Schönberg and Alexandra Spitz-Oener (2014). From sick man of Europe to economic superstar: Germany's resurgent economy. *Journal of Economic Perspectives* 28 (1), pages 167–88.

Dustmann, Christian, Bernd Fitzenberger and Markus Zimmermann (2022). Housing expenditure and income inequality. *The Economic Journal* 132 (645), pages 1709–1736.

Dustmann, Christian, Johannes Ludsteck and Uta Schönberg (2009). Revisiting the German wage structure. *Quarterly Journal of Economics* 124 (2), pages 843–881.

D'Haultfœuille, Xavier, Arnaud Maurel and Yichong Zhang (2018). Extremal quantile regressions for selection models and the black–white wage gap. *Journal of Econometrics* 203 (1), pages 129–142.

Ehrich, Malte, Abdul Musasib and Devesh Roy (2018). The Hartz reforms and the German labor force. *European Journal of Political Economy* 55, pages 284–300.

Elass, Kenza (2024). Male and female selection effects on gender wage gaps in three countries. *Labour Economics* 87, page 102506.

Eloundou, Tyna, Sam Manning, Pamela Mishkin and Daniel Rock (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science* 384 (6702), pages 1306–1308.

Ermisch, John F. and Robert E. Wright (1994). Interpretation of negative sample selection effects in wage offer equations. *Applied Economics Letters* 1 (11), pages 187–189.

Faggian, Alessandra and Rachel S. Franklin (2014). Human capital redistribution in the USA: The migration of the college-bound. *Spatial Economic Analysis* 9 (4), pages 376–395.

Felbermayr, Gabriel, Michele Battisti and Sybille Lehwald (2016). Einkommensungleichheit in Deutschland, Teil 1: Gibt es eine Trendumkehr? *ifo Schnelldienst* 69 (13), pages 28–37.

Fernandez-Val, Aico Ivan van Vuuren and Francis Vella (2024a). Hours worked and the US distribution of real annual earnings 1976-2019. *Journal of Applied Econometrics* 39, pages 659–678.

Fernandez-Val, Aico Ivan van Vuuren and Francis Vella (2024b). Nonseparable sample selection models with censored selection rules. *Journal of Applied Econometrics* 240, page 105088.

Fernandez-Val, Ivan, Aico van Vuuren, Francis Vella and Franco Peracchi (2023). Selection and the distribution of female real hourly wages in the United States. *Quantitative Economics* 14 (2), pages 571–607.

Fitzenberger, Bernd and Jakob de Lazzer (2022). Changing selection into full-time work and its effect on wage inequality in Germany. *Empirical Economics* 62, pages 247–277.

Fitzenberger, Bernd and Karsten Kohn (2006). Skill wage premia, employment, and cohort effects: Are workers in Germany all of the same type? *IZA Discussion Paper* 2185.

Fitzenberger, Bernd and Arnim Seidlitz (2020). The 2011 break in the part-time indicator and the evolution of wage inequality in Germany. *Journal for Labour Market Research* 54, pages 1–14.

Flossmann, Anton L. and Winfried Pohlmeier (2006). Causal returns to education: a survey on empirical evidence for Germany. *Jahrbücher für Nationalökonomie und Statistik* 226 (1), pages 6–23.

Fortin, Nicole, Thomas Lemieux and Sergio Firpo (2011). Decomposition methods in economics. In *Handbook of Labor Economics*, Volume 4A, pages 1–102. Elsevier.

Freeman, Richard B. (1999). The economics of crime. In *Handbook of Labor Economics*, Volume 3, pages 3529–3571. Elsevier.

Fuchs-Schündeln, Nicola, Dirk Krueger and Mathias Sommer (2010). Inequality trends for Germany in the last two decades: A tale of two countries. *Review of Economic Dynamics* 13 (1), pages 103–132.

Gallego-Granados, Patricia (2019). The part-time wage gap across the wage distribution. *DIW Discussion Paper* 1791.

Gallego-Granados, Patricia and Katharina Wrohlich (2019). Selection into employment and the gender wage gap across the distribution and over time. *IZA Discussion Paper* 12859.

Gartner, Hermann (2005). The imputation of wages above the contribution limit with the German IAB employment sample. *FDZ Methodenreport* 2(2005), pages 1–8.

Gebel, Michael and Friedhelm Pfeiffer (2010). Educational expansion and its heterogeneous returns for wage workers. *Journal of Contextual Economics–Schmollers Jahrbuch* 130 (1), pages 19–42.

Geyer, Johannes, Peter Haan and Katharina Wrohlich (2015). The effects of family policy on maternal labor supply: Combining evidence from a structural model and a quasi-experimental approach. *Labour Economics* 36, pages 84–98.

Glitz, Albrecht and Daniel Wissmann (2021). Skill premiums and the supply of young workers in Germany. *Labour Economics* 72, page 102034.

Goebel, Jan, Markus M. Grabka, Stefan Liebig, Martin Kroh, David Richter, Carsten Schröder and Jürgen Schupp (2019). The German Socio-Economic Panel (SOEP). *Jahrbücher für Nationalökonomie und Statistik* 239 (2), pages 345–360.

Goldin, Claudia (2006). The quiet revolution that transformed women's employment, education, and family. *American Economic Review* 96 (2), pages 1–21.

Goldin, Claudia (2014). A grand gender convergence: Its last chapter. *American Economic Review* 104 (4), pages 1091–1119.

Goldin, Claudia (2021). *Career and family: Women's century-long journey toward equity.* Princeton University Press.

Goller, Daniel, Christian Gschwendt and Stefan C. Wolter (2025). This time it's different – Generative artificial intelligence and occupational choice. *Labour Economics*, page 102746.

Gould, Eric D. and Alexander Hijzen (2016). Growing apart, losing trust? The impact of inequality on social capital. *IMF Working Paper* 16/176.

Grabka, Markus (2017). SOEP 2016 - Codebook for the PEQUIV File 1984-2016: CNEF Variables with Extended Income Information for the SOEP. *SOEP Survey Papers, Series D - Variable Descriptions and Coding.*

Green, Andrew (2024). Artificial intelligence and the changing demand for skills in the labour market. *OECD Artificial Intelligence Papers* 14.

Grewenig, Elisabeth (2021). School track decisions and teacher recommendations: Evidence from German state reforms. *ifo Working Paper* 353.

Hartung, Benjamin, Philip Jung and Moritz Kuhn (2018). What hides behind the German labor market miracle? Unemployment insurance reforms and labor market dynamics. *CEPR Discussion Paper* 133280.

Heckman, James J. (1974). Shadow prices, market wages, and labor supply. *Econometrica* 42 (4), pages 679–694.

Heckman, James J. (1979). Sample selection bias as a specification error. *Econometrica* 47 (1), pages 153–161.

Heckman, James J. and Tim Kautz (2012). Hard evidence on soft skills. *Labour Economics* 19 (4), pages 451–464.

Henderson, Michael B., Philipp Lergetporer, Paul E. Peterson, Katharina Werner, Martin R. West and Ludger Woessmann (2015). Is seeing believing? How Americans and Germans think about their schools. *ifo Working Paper* 202.

Hochmuth, Brigitte, Britta Kohlbrecher, Christian Merkl and Hermann Gartner (2021). Hartz IV and the decline of German unemployment: A macroeconomic evaluation. *Journal of Economic Dynamics and Control* 127, page 104114.

Hoffmann, Florian and Thomas Lemieux (2016). Unemployment in the Great Recession: a comparison of Germany, Canada, and the United States. *Journal of Labor Economics* 34 (S1), S95–S139.

Huber, Martin and Giovanni Mellace (2014). Testing exclusion restrictions and additive separability in sample selection models. *Empirical Economics* 47, pages 75–92.

Huber, Martin and Blaise Melly (2015). A test of the conditional independence assumption in sample selection models. *Journal of Applied Econometrics* 30 (7), pages 1144–1168.

Hutter, Christian, Francesco Carbonero, Sabine Klinger, Carsten Trenkler and Enzo Weber (2022). Which factors were behind Germany's labour market upswing? A data-driven approach. *Oxford Bulletin of Economics and Statistics* 84 (5), pages 1052–1076.

Hyslop, Dean R. and David C. Maré (2005). Understanding New Zealand's changing income distribution, 1983–1998: A semi-parametric analysis. *Economica* 72 (287), pages 469–495.

Jenkins, Stephen P. (2024). Getting the measure of inequality. *Oxford Open Economics* 3, pages 156–166.

Jerrim, John and Lindsey Macmillan (2015). Income inequality, intergenerational mobility, and the Great Gatsby Curve: Is education the key? *Social Forces* 94 (2), pages 505–533.

Jessen, Robin (2019). Why has income inequality in Germany increased from 2002 to 2011? A behavioral microsimulation decomposition. *Review of Income and Wealth* 65 (3), pages 540–560.

Juhn, Chinhui, Dae Il Kim and Francis Vella (2005). The expansion of college education in the United States: Is there evidence of declining cohort quality? *Economic Inquiry* 43 (2), pages 303–315.

Katz, Lawrence F. and David H. Autor (1999). Changes in the wage structure and earnings inequality. In *Handbook of Labor Economics*, Volume 3, pages 1463–1555. Elsevier.

Katz, Lawrence F. and Kevin M. Murphy (1992). Changes in relative wages, 1963–1987: supply and demand factors. *The Quarterly Journal of Economics* 107 (1), pages 35–78.

Kelly, Morgan (2000). Inequality and crime. *Review of Economics and Statistics* 82 (4), pages 530–539.

Kügler, Alice, Uta Schönberg and Ranghild Schreiner (2018). Productivity growth, wage growth and unions. In *ECB Conference Proceedings: Price and wage-setting in advanced economies*, pages 215–247.

Kuhn, Andreas (2019). The subversive nature of inequality: Subjective inequality perceptions and attitudes to social inequality. *European Journal of Political Economy* 59 (5), pages 331–344.

Kunze, Astrid (2018). The gender wage gap in developed countries. *The Oxford Handbook of Women and the Economy*, pages 369–394.

Launov, Andrey and Klaus Wälde (2013). Estimating incentive and welfare effects of nonstationary unemployment benefits. *International Economic Review* 54 (4), pages 1159–1198.

Launov, Andrey and Klaus Wälde (2016). The employment effect of reforming a public employment agency. *European Economic Review* 84, pages 140–164.

Lemieux, Thomas (2006a). Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *American Economic Review* 96 (3), pages 461–498.

Lemieux, Thomas (2006b). Postsecondary education and increasing wage inequality. *American Economic Review* 96 (2), pages 195–199.

Luthra, Renee Reichl and Jennifer Flashman (2017). Who benefits most from a university degree?: A cross-national comparison of selection and wage returns in the US, UK, and Germany. *Research in Higher Education* 58 (8), pages 843–878.

Maasoumi, Esfandiar and Le Wang (2019). The gender gap between earnings distributions. *Journal of Political Economy* 127, pages 2438–2504.

Maaz, Kai, Marko Neumann, Ulrich Trautwein, Wolfgang Wendt, Rainer Lehmann and Jürgen Baumert (2008). Der Übergang von der Grundschule in die weiterführende Schule: Die Rolle von Schüler-und Klassenmerkmalen beim Einschätzen der individuellen Lernkompetenz durch die Lehrkräfte. *Swiss Journal of Educational Research* 30 (3), pages 519–548.

Manning, Alan and Barbara Petrongolo (2008). The part-time penalty for women in Britain. *Economic Journal* 118 (526), F28–F51.

Martins, Pedro S. and Pedro T. Pereira (2004). Does education reduce wage inequality? Quantile regression evidence from 16 countries. *Labour Economics* 11 (3), pages 355–371.

Matteazzi, Eleonora, Ariane Pailhe and Anne Solaz (2014). Part-time wage penalties for women in prime age. *Industrial and Labor Relations Review* 67 (3), pages 955–985.

Melly, Blaise and Giulia Santangelo (2014). The evolution of the gender wage gap: 1968-2008. *Working Paper*.

Möller, Joachim (2016). Lohnungleichheit – gibt es eine Trendwende? *IAB-Discussion Paper* 9/2016.

Müller-Benedict, Volker and Thomas Gaens (2020). A new explanation for grade inflation: The long-term development of German university grades. *European Journal of Higher Education* 10 (2), pages 181–201.

Mulligan, Casey B. and Yona Rubinstein (2008). Selection, investment, and women's relative wages over time. *The Quarterly Journal of Economics* 123 (3), pages 1061–1110.

NEPS Network (2022). National Educational Panel Study, Scientific Use File of Starting Cohort Adults. Leibniz Institute for Educational Trajectories (LIfBi), Bamberg.

Oaxaca, Ronald (1973). Male-female wage differentials in urban labor markets. *International Economic Review* 14 (3), pages 693–709.

OECD (2021). *Bildung auf einen Blick 2021*. `https://www.oecd-ilibrary.org/content/publication/6001821ow`. Accessed: 11-17-2023.

Olivetti, Claudia, Jessica Pan and Barbara Petrongolo (2024). The evolution of gender in the labor market. In *Handbook of Labor Economics*, Volume 5, pages 619–677. Elsevier.

Olivetti, Claudia and Barbara Petrongolo (2008). Unequal pay or unequal employment? A cross-country analysis of gender gaps. *Journal of Labor Economics* 26 (4), pages 621–654.

Ordemann, Jessica and Friedhelm Pfeiffer (2022). The evolution of educational wage differentials for women and men in Germany, from 1996 to 2019. *Journal for Labour Market Research* 56 (17), pages 1–12.

Peichl, Andreas, Paul Hufe and Marc Stöckli (2018). Ökonomische Ungleichheit in Deutschland – ein Überblick. *Perspektiven der Wirtschaftspolitik* 19 (3), pages 185–199.

Pereda-Fernández, Santiago (2025). Decomposition of differences in distribution under sample selection and the gender wage gap. *Journal of Business & Economic Statistics* 43 (2), pages 378–390.

Picchio, Matteo and Chiara Mussida (2011). Gender wage gap: A semi-parametric approach with sample selection correction. *Labour Economics* 18 (5), pages 564–578.

Pickett, Kate E. and Richard G. Wilkinson (2015). Income inequality and health: a causal review. *Social Science & Medicine* 128, pages 316–326.

Raaum, Oddbjørn and Knut Røed (2006). Do business cycle conditions at the time of labor market entry affect future employment prospects? *The Review of Economics and Statistics* 88 (2), pages 193–210.

Reinhold, Mario and Stephan Thomsen (2017). The changing situation of labor market entrants in Germany. *Journal for Labour Market Research* 50 (1), pages 161–174.

Riphahn, Regina T. and Rebecca Schrader (2020). Institutional reforms of 2006 and the dramatic rise in old-age employment in Germany. *Industrial Labor Relations Review* 20 (5), pages 1185–1225.

Riphahn, Regina T. and Michael Zibrowius (2016). Apprenticeship, vocational training, and early labor market outcomes – evidence from East and West Germany. *Education Economics* 24 (1), pages 33–57.

Rothe, Thomas and Klaus Wälde (2017). Where did all the unemployed go? Non-standard work in Germany after the Hartz reforms. *IAB Discussion Paper* 1817.

Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung (2019). Aufstiegschancen sichern, Arbeitsanreize stärken, Jahresgutachten 2019/20, pages 316–374.

Sauer, Stefan and Timo Wollmershäuser (2021). Fachkräftemangel wird zunehmend zur Belastung für die deutsche Wirtschaft. *ifo Schnelldienst Digital* 2 (17), pages 1–4.

Sologon, Denisa M., Vanda Almeida and Philippe Van Kerm (2019). Accounting for the distributional effects of the 2007-2008 crisis and the economic adjustment program in Portugal. *Luxembourg Institute of Socio-Economic Research Discussion Paper* 2019-05.

Statistisches Bundesamt (1976). Statistisches Jahrbuch für die Bundesrepublik Deutschland, 1976.

Statistisches Bundesamt (2019). Statistisches Jahrbuch für die Bundesrepublik Deutschland, 2019.

Statistisches Bundesamt (2024). *Teilzeitquote erneut leicht gestiegen auf 31% im Jahr 2023*. Statistisches Bundesamt. URL: https://www.destatis.de/DE/Presse/Pressemitteilungen/2024/04/PD24_N017_13.html (visited on 17/10/2024).

von Wachter, Till (2020). The persistent effects of initial labor market conditions for young adults and their sources. *Journal of Economic Perspectives* 34 (4), pages 168–194.

Walker, Ian and Yu Zhu (2008). The college wage premium and the expansion of higher education in the UK. *The Scandinavian Journal of Economics* 110 (4), pages 695–709.

Wolter, Andrä and Christian Kerst (2015). The 'academization' of the German qualification system: Recent developments in the relationships between vocational training and higher education in Germany. *Research in Comparative and International Education* 10 (4), pages 510–524.

Zucman, Gabriel (2019). Global wealth inequality. *Annual Review of Economics* 11 (1), pages 109–138.

Zühlke, Anne, Philipp Kugler, Armin Hackenberger and Tobias Brändle (2022). Accounting for dropout risk and upgrading in educational choices: New evidence for lifetime returns in Germany. *Education Economics* 30 (6), pages 1–16.