EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

University of Tübingen
Working Papers in
Economics and Finance

No. 118

The pure effect of social preferences on regional location choices: The evolving dynamics of convergence to a steady state population distribution

by

Oded Stark, Wiktor Budzinski, Grzegorz Kosiorowski

Faculty of Economics and Social Sciences
www.wiwi.uni-tuebingen.de

The pure effect of social preferences on regional location choices: The evolving dynamics of convergence to a steady state population distribution

Oded Stark

Universities of Bonn, Tuebingen, and Warsaw. Address: ZEF, University of Bonn, Walter-Flex-Strasse 3, D-53113 Bonn, Germany. E-mail: ostark@uni-bonn.de

Wiktor Budzinski

University of Warsaw. E-mail: wbudzinski@wne.uw.edu.pl

Grzegorz Kosiorowski

Cracow University of Economics. E-mail: kosiorog@uek.krakow.pl

Abstract

This paper tracks the consequences of individuals' desire to align their location with their social preferences. The social preference studied in the paper is distaste for relative deprivation, measured in a cardinal manner. Location is conceived as social space, with individuals choosing to relocate if, as a result, their relative deprivation will be reduced, holding their incomes constant. Conditions are provided under which the associated dynamics reaches a spatial steady state, the number of periods it takes to reach a steady state is specified, and light is shed on the robustness of the steady state outcome. By way of simulation it is shown that for large populations, a steady state of the relocation dynamics is almost always reached, typically in one period, and that cycles are more likely to occur when the populations' income distributions are more equal.

Keywords

Social preferences; Distaste for low relative income; A cardinal measure of income relative deprivation; Interregional locational choices; Relocation dynamics; Steady-state spatial distribution

JEL Classification

C62, C63, R12, R13, Z13

1. INTRODUCTION

By now there is widespread recognition, based on mounting evidence, that comparisons with others impinge significantly on wellbeing, and elicit substantial behavioral responses. The received literature reveals that the comparisons which matter for an individual's sense of wellbeing are those made by looking "up" the income hierarchy, rather than by looking "down." A large literature that supports the "upward comparison" hypothesis is reviewed in Frey and Stutzer (2002), Walker and Smith (2002), and Stark (2013), for example. Engaging in interpersonal comparisons affects the individuals' sense of wellbeing and influences their behavior, including in relation to where to locate. Yet there has been no systematic inquiry into how the *pure* effect of social comparisons determines locational outcomes. This paper takes a step towards filling this lacuna.

The paper characterizes the steady state distribution of a population of n individuals who are homogeneous in preferences and heterogeneous in incomes. The individuals who to begin with are in region A can relocate at no cost to themselves between region A and region B. We make two main assumptions: that the individuals exhibit strong social preferences, and that their incomes are held constant. The reason for making the first assumption is given in the preceding paragraph. The reason for making the second assumption is to allow us to concentrate on essentials, namely to facilitate a study of the pure effect of location-specific dissatisfaction that arises from falling behind others in the income distribution. Social preferences take the form of distaste for falling behind others with respect to income; in other words, social preferences represent the negative influence of unfavorable income comparisons on the individuals' sense of wellbeing.

We model social preferences as distaste for relative deprivation (defined later on in this paragraph). Because incomes are held constant, the wellbeing of an individual is solely a function of the extent to which that individual's location aligns with his social preferences. To begin with we assume that across the income hierarchy, the income differences between any two adjacent individuals are the same. We obtain three interesting results. First, the process of relocation reaches a spatial steady state (namely the movement between locations ceases, with no individual being able to improve his wellbeing by engaging in further movement). Second, under relative deprivation, the steady state outcome is a sharp bifurcation, with the individual whose income is the highest staying in region A, and all the other individuals relocating to and staying in region B. Third, regardless of how relative deprivation is measured, whether as the aggregate of the income excesses divided by the size of the population or as the distance from below the mean income, the result is a spatial steady state. However, when incomes differ but are not equally spaced, a spatial steady state distribution may or may not be achieved. We show that for $n=3$ (and trivially for $n=2$) a spatial steady state distribution will always be achieved. When $n=4$ and $n=5$, we specify conditions under which a spatial steady state will not be reached, and complementary conditions under which the process of relocation reaches a spatial steady state. We comment on the difficulty of obtaining predictions for the case of $n \geq 6$, and we then resort to a simulation procedure that enables us to gain insights from analyzing this case. Quite remarkably, we find that the outcome to which the simplified model (where the income differences between any two adjacent individuals are the same) gives rise is a generic outcome of the dynamics of locational choices for a large number of individuals with an arbitrary distribution of incomes. When the size of the population increases, the probability of reaching a spatial steady state, as well as the average number of periods needed to reach that state, tend to 1. We also find that regardless of the size of the population, a spatial steady state is more likely to be reached in the case of populations with less equal income distributions.

Considerable empirical evidence finds that relative deprivation is a statistically significant explanatory variable of a notable case of locational moves, namely of migration behavior. Stark and Taylor (1991) show that relative deprivation increases the probability that the labor time of household members will migrate from rural Mexico to the US to work. The significance of relative deprivation as an explanatory variable of labor migration received additional support in several more recent studies. Quinn (2006) reports that relative deprivation is a significant motivating factor in domestic migration decisions in Mexico. Stark et al. (2009) explore the relationship between aggregate relative poverty, which is functionally related to aggregate relative deprivation, and migration. Drawing on Polish regional data, they demonstrate that migration from a region is positively correlated with the aggregate relative deprivation in the region. Czaika (2012) finds that, in India, relative deprivation is an important factor in deciding whether a household member should migrate, especially for migration over a short distance. Basarir (2012) observes that people in Indonesia are willing to bear a loss of absolute wealth if there is a relative wealth gain from migration. Jagger et al. (2012) report that relative deprivation is a significant explanatory variable of circular migration in Uganda. Vernazza (2013) concludes that, even though interstate migration in the US confers substantial increases in absolute income, the trigger for migration is relative deprivation (low relative income), not low absolute income. Drawing on data from the 2000 US census, Flippen (2013) shows that both blacks and whites who migrate from the North to the South generally have average lower absolute incomes than their stationary northern peers, yet enjoy significantly lower relative deprivation, and that the relative deprivation gains for blacks are substantially larger than those for whites. Hyll and Schneider (2014) use a data set collected in the German Democratic Republic in 1990 to show that aversion to relative deprivation enhanced the propensity to migrate to western Germany. Kafle et al. (2018) use comparable longitudinal data from integrated household and agriculture surveys from Tanzania, Ethiopia, Malawi, Nigeria, and Uganda, and find that wealth relative deprivation is positively associated with migration.

The remainder of this paper is organized as follows. In Section 2 we present the assumptions of our analytical framework. In Section 3 we analyze a model of choice of location between two regions under the assumption that the income differences between any two adjacent individuals are equal to 1. In Sections 4 and 5 we analyze the consequences of relaxing several of the model's assumptions: in Section 4 we relax the assumptions about the number of individuals of each income, and about the size of the equal income difference between adjacent individuals being equal to 1. We find that the results of Section 3 are not contingent on these assumptions. In Section 5 we revoke the assumption that the income differences between two adjacent individuals are the same. We establish conditions for the existence of a spatial steady state of the location dynamics in this case. We accomplish this analytically for populations of size $n \leq 5$, and for larger populations - by means of simulation. Proofs of the claims made in Section 5 are in Appendix A. In Section 6 we comment on an adjustment of the models to the possibility that incomes can change. In Section 7 we discuss the possibilities that to begin with not all the individuals might be in region A, and that there might be more than one new region available for the individuals to move to. In Section 8 we conclude.

2. CHARACTERIZING THE INDIVIDUALS

Let there be a population of n individuals, where n is a natural number. The income of individual i is i, $i = 1, 2, ..., n$ (namely the individual's income is the individual's name). To begin with all the individuals are in region A. Let (empty) region B come into being or become accessible such that moving between the two regions is possible, and is cost free. In all relevant respects, the two regions are identical. This implies that there is no reason, arising from a difference in the regions' amenities, for an individual to prefer one region to the other. The individuals want to be in the region that better aligns with their social preferences. When, in terms of the outcome of social comparisons, the regions are equally attractive (a tie), the individuals do not relocate. Once the individuals are in a region, the region becomes instantaneously their exclusive sphere of comparison. However, in response to the actual distribution of people between the two regions, the individuals can relocate as many times as they wish, at no cost to themselves. Put differently, the individuals base their location decisions on the observed current state, without simultaneously forming expectations how other individuals will behave. For ease of exposition, we refer to the steps in the process of selecting location as periods, with the initial period being referred to as zero.

3. MEASURES OF SOCIAL PREFERENCES

3.1 Indices of relative deprivation (RD)

3.1.1 Relative deprivation measured as the aggregate of the income excesses divided by the size of the population

Let $(x_1, x_2, \ldots, x_n)$ be an ordered vector of incomes of a given population of n individuals, namely x_i is the income of individual i, and $x_1 \le x_2 \le \ldots \le x_n$. Then, we measure the relative deprivation of individual i as follows.

Definition 1: $RD(i) \equiv \frac{1}{n}\sum_{k=i+1}^{n} \max\{x_k - x_i, 0\}$ for $i = 1, 2, ..., n-1$; $RD(n) \equiv 0$.

A rationale underlying this measure is provided in Appendix B. Under the assumption of Section 2 that $x_i = i$, we obtain that

Definition 2: $RD(i) \equiv \frac{1}{n}\sum_{k=i+1}^{n} \max\{k - i, 0\}$ for $i = 1, 2, ..., n-1$; $RD(n) \equiv 0$.

To begin with in period zero the *n* individuals are in region A. In the subsequent period, all the individuals who experience relative deprivation and believe that they will experience none upon relocating to region B move to region B. Namely:

A	B
n	
	$n-1$
	$n-2$
	.
	.
	.
	1

Claim 1: Under relative deprivation measured as per Definition 2, the division in which *n* is in region A and the remainder of the population is in region B constitutes the spatial steady state distribution.

Proof: We consider individual k , $k = 1, 2, ..., n-1$ who in period 1 weighs whether to stay in region B or whether to move back to region A. If he stays in region B:

$$RD(k)\big|_{k\in B} = \frac{1}{n-1}[(k+1)-k+(k+2)-k+...+(n-1)-k] = \frac{n-k}{2}\cdot\frac{n-k-1}{n-1}.$$

If he were to return to region A:

$$RD(k)\big|_{k\in A} = \frac{n-k}{2}.$$

Because $\frac{n-k-1}{n-1} < 1$, individual k will prefer to stay in region B. And because this holds true for any $k = 1, 2, ..., n-1$, none of the *n* individuals will have an incentive to relocate and, thus, the observed state, as depicted in the box diagram above, is the spatial steady state. Q.E.D.

3.1.2 Relative deprivation measured as the distance from below the mean income

The relative deprivation of an individual can also be measured by how much the individual needs to increase his income in order to obtain the average income of the region in which he is located.

Definition 3: $RD(i) \equiv \max\{\bar{x} - x_i, 0\}$ where $\bar{x}$ is the average income in the region in which individual *i* is located.

Under the assumption of Section 2 that $x_i = i$, we formulate the following definition.

Definition 4: $RD(i) \equiv \max\{\bar{x} - i, 0\}$.

We show that the dynamics of movement between the two regions driven by relative deprivation, measured as per Definition 3, differs only slightly from the dynamics of movement driven by relative deprivation measured as per Definition 1.

Claim 2: Under relative deprivation measured as per Definition 4, the division in which *n* is in region A and the remainder of the population is in region B constitutes the spatial steady state distribution.

Proof: To begin with in period zero the *n* individuals are in region A. In the subsequent period, all the individuals who are relatively deprived - in this case, the individuals whose incomes are lower than the average income in region A - will move to region B, while the other individuals will remain in region A. Thus, individuals $n, n-1, ..., m$ where $m = \frac{n}{2} + 1$ if n is even, and individuals $n, n-1, ..., m$ where $m = \frac{n+1}{2}$ if n is odd, will remain in region A, whereas individuals $m-1, ..., 2, 1$ will move to region B. But now the average income in region A becomes higher, so in the subsequent period the individuals whose income is below the average income of those remaining in region A become relatively deprived and they will, thus, be better off moving to region B. This process will continue until only individual *n* remains in region A.

We note that none of the individuals who have relocated to region B will find it attractive to return to region A even after the subsequent arrivals in region B of the higher income individuals. Thus, again, a spatial distribution such that individual *n* is in region A while individuals $1, 2, ..., n-1$ are in region B constitutes the steady state spatial distribution. To see this, consider individual k, $k = 1, 2, ..., n-1$. The average income in region B in the "alleged" steady state distribution is $\frac{n}{2}$, and this is lower than $\frac{n+k}{2}$, the average income that individual k will experience if he were to return to region A. Thus, if $k \geq \frac{n}{2}$, then individual k does not have an incentive to move back to region A because he is not relatively deprived in region B. And if $k < \frac{n}{2}$, namely if individual k is relatively deprived in region B, then his relative deprivation there is $\frac{n}{2} - k$, and this is lower than his relative deprivation will be in region A, which is $\frac{n+k}{2} - k$. Hence, no further movement between the regions will occur. Q.E.D.

Comment: in this case, reaching the spatial steady state will take $\lfloor \log_2(n-1) \rfloor + 1$ periods, where the symbol $\lfloor x \rfloor$ denotes the biggest integer that is not greater than x. For example, when $n = 8$, the number of periods it takes to reach the steady state will be $\lfloor \log_2(8-1) \rfloor + 1 = 2 + 1 = 3$.

From now on, unless explicitly stated otherwise, we use Definition 1 as our "default" measure of the relative deprivation of individual i.

4. ROBUSTNESS

Claim 3: Having more than one individual of each income does not change the spatial steady state distribution.

Proof: We assume that there are l individuals of each income (where l is a natural number). This means that the income of individuals $1, 2, \ldots, l$ is 1, the income of individuals $l+1, l+2, \ldots, 2l$ is 2 and, in general, the income of individuals $(k-1)l+1, (k-1)l+2, \ldots, kl$ is k for $k = 1, 2, ..., n$. Take the case of Section 3.1.1. We have l individuals whose income is n staying in region A, and the rest of the individuals moving to region B. It is easy to verify that for $k = 1, 2, ..., n-1$ and any $\tilde{k}$ such that $(k-1)l < \tilde{k} \leq kl$ (meaning that we consider individual $\tilde{k}$ whose income is k), $RD(\tilde{k})\big|_{\tilde{k} \in B} = \frac{n-k}{2} \cdot \frac{n-k-1}{n-1}$, namely that the relative deprivation experienced by any individual whose income is k, $k = 1, 2, ..., n-1$, is the same as the relative deprivation already calculated in Section 3.1.1. If one of the l individuals whose income is k (namely individual $\tilde{k}$ such that $(k-1)l < \tilde{k} \leq kl$) were to move back to region A, then his relative deprivation there will be:

$$RD(\tilde{k})\big|_{\tilde{k} \in A} = \frac{l}{l+1}(n-k) .$$

Because l is a natural number, $\frac{l}{l+1}(n-k) \geq \frac{n-k}{2} > \frac{n-k}{2} \cdot \frac{n-k-1}{n-1}$, so staying in region B will be preferable to returning to region A. By similar reasoning it follows that having l individuals of each income does not change the results obtained for different indices of social preferences: the mean income defined in Section 3.1.2 does not change in such a setting and, therefore, the location decisions of the individuals will be the same. Q.E.D.

Claim 4: Affine transformation of the vector of incomes does not change the spatial steady state distribution.

Proof: Instead of the vector of incomes $(1, 2, \ldots, n)$ we consider the vector $(\alpha \cdot 1 + \beta, \alpha \cdot 2 + \beta, \ldots, \alpha \cdot n + \beta)$, with $\alpha > 0$ and $\beta > -\alpha$. Relative deprivation can be viewed as a function of the incomes of all the individuals.[1] Then, we can see that the relative deprivation function is homogeneous of degree one, namely that the relative deprivation of individual k, which here and only here we now denote as RD_k, observes

$RD_k\left(\alpha\cdot 1+\beta,\alpha\cdot 2+\beta,\ldots,\alpha\cdot n+\beta\right)=\alpha RD_k\left(1,2,\ldots,n\right)$. Therefore, when comparing the relative deprivation for a given individual between the two regions, we note that an affine transformation of all the incomes results in rescaling relative deprivation by the same factor in both regions which, thus, does not change the obtained results. Q.E.D.

5. GENERALIZATION: THE CASE OF ARBITRARY INCOME DIFFERENCES

We now ask about the consequences of revoking the assumption that the income differences between all pairs of adjacent individuals are the same. When incomes are not equally spaced, the process of selection of location may or may not reach a spatial steady state.

Consider the following two examples. First, suppose that there are four individuals with incomes 12, 11, 8, and 5 who to begin with are all in region A. Let empty region B come into being or become accessible. The evolving dynamics is depicted by the following sequence:

A	B
12	
11	
8	
5	

→

A	B
12	
	11
	8
	5

Because none of the individuals now has an incentive to move, we conclude that a spatial steady state is reached in just one period, with individual 12 in region A, and individuals 11, 8, and 5 in region B.

Second, suppose that the income of the poorest individual is 1 rather than 5, so that we now have four individuals with incomes 12, 11, 8, and 1. Such a change alters the calculus as reported above of the lowest income individual and influences his region of choice which, in turn, affects the calculus of individual 8 and his location decision; the change in the income of the lowest income individual inflicts a “location externality” on the second lowest income individual. To see this, let all four individuals again begin in region A. Now empty region B comes into being. The evolving dynamics is depicted by the following sequence:

A	B
12	
11	
8	
1	

→

A	B
12	
	11
	8
	1

→

A	B
12	
	11
	8
1	

→

A	B
12	
	11
8	
1	

→

A	B
12	
	11
8	
	1

→

A	B
12	
	11
	8
	1

→

A	B
12	
	11
	8
1	

→ . . .

We see that in this case the process repeats itself ad infinitum, and a steady state is not reached. The perpetual movement in this example (in which individual 8 will always want to be located where individual 1 is located, and individual 1 will always want to be located where individual 8 is absent) emanates from the fact that the behavior of individual 8 is "tied" to the presence of individual 1 in that this presence reduces the agony from looking up at individual 11 or at individual 12.

This second example can be generalized. We formulate conditions under which the process of selecting a location will reach or fail to reach a spatial steady state by stating and proving three Lemmas.

Let the income of individual i be x_i, $i = 1, 2, ..., n$, and let $x_1 < x_2 < \ldots < x_n$. Assume that the social preferences of the individuals are measured by their relative deprivation defined as the aggregate of income excesses divided by the size of the population. Because an analysis based on relative deprivation defined as the distance from below the mean income is analogous, it will be skipped.[2] Then:

Lemma 1. When $n = 3$, a spatial steady state will always be reached: individual 3 will be in region A, and individuals 2 and 1 will be in region B.

Lemma 2. When $n = 4$, the distribution of the individuals between the two regions will perpetually change and a spatial steady state will not be reached iff $2x_4 + x_2 < 3x_3$ and $3x_4 + x_1 < 2(x_3 + x_2)$. Otherwise, a spatial steady state will be reached.

Lemma 3. When $n = 5$, the distribution of the individuals between the two regions will perpetually change and a steady state will not be obtained iff $2x_5 + x_1 < x_4 + x_3 + x_2$ and $x_5 + x_2 < x_4 + x_3$ and $x_5 + x_3 < 2x_4$. Otherwise, a spatial steady state will be reached.

The proofs of the lemmas are tedious, and are thus relegated to Appendix A.

Although it would be possible to construct similar criteria for any $n > 5$, the respective formulas become increasingly longer and more complicated when n increases beyond 5. In particular, as shown in Appendix A, the three lemmas are proved by considering the step-by-step behavior of the individuals in each period. For $n \leq 5$, we establish that individuals move in only one direction in a period. This may not be the case, however, when $n \geq 6$. Then, we can have two individuals moving in opposite directions in the very same period. To see this, consider an example of six individuals with incomes 5.3, 5.2, 5.1, 5, 4, and 1. In period one, as usual, everyone except the highest income individual 5.3 will migrate to region B. Then, in period two, as can be easily checked, only the two individuals with the lowest incomes will have an incentive to move back to region A, which they do. Hence, in period three, individuals 5.3, 4, and 1 are in region A, and the other individuals, namely individuals 5.2, 5.1, and 5 are in region B. In this setting, only individual 5 will want to move in period three, which leads to the following distribution of the individuals:

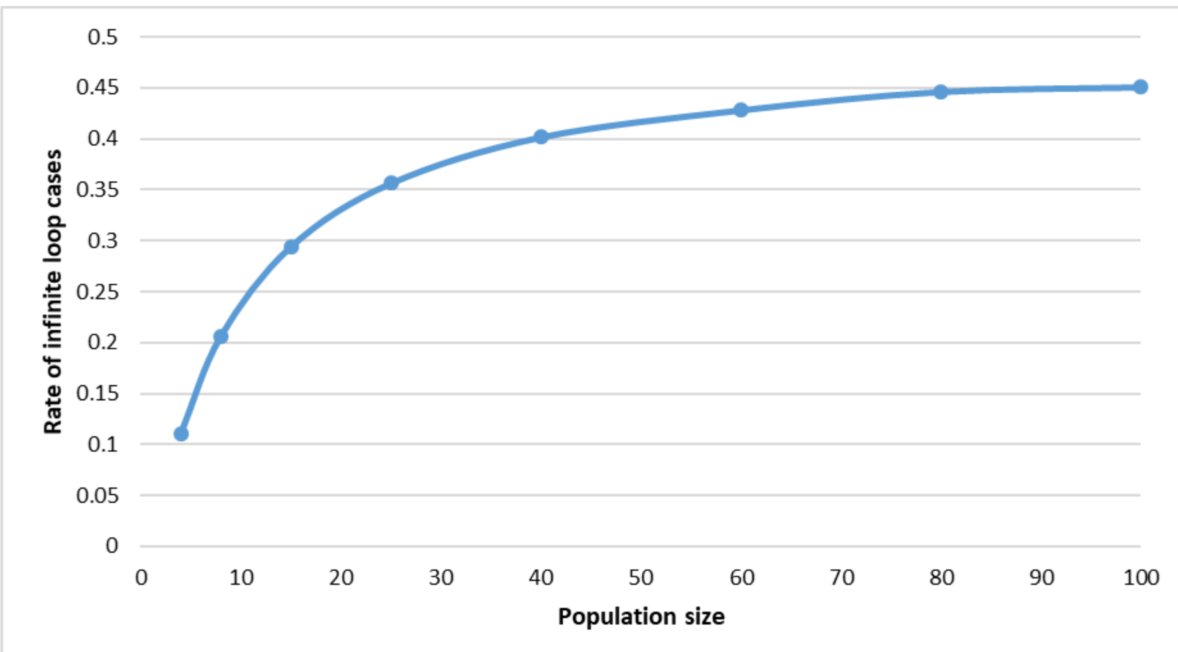
A	B
5.3	
	5.2
	5.1
5	
4	
1	

Now both individuals 1 and 5.1 will want to change their location: the relative deprivation of individual 1 is $\frac{11.3}{4} = \frac{33.9}{12}$ in region A, which is higher than $\frac{8.3}{3} = \frac{33.2}{12}$, his relative deprivation if he were to locate to region B. (For ease of reference, we mix decimal notation with fraction notation, the non-elegance of such a blending notwithstanding.) Analogously, the relative deprivation of individual 5.1 is $\frac{0.1}{2}=\frac{0.5}{10}$ in region B, which is higher than $\frac{0.2}{5}=\frac{0.4}{10}$, his relative deprivation if he were to locate in region A. Because both these individuals will indeed move, we have a simultaneous two-way movement.

In sum, in order to determine the outcome of the relocation dynamics for $n \geq 6$, we will need to distinguish in each period not only between reaching a steady state and a move of exactly one individual, but also between patterns of behavior that involve either a one-way or a simultaneous two-way movement of many individuals. For $n \leq 5$, this distinction could be obtained by means of a single inequality, whereas for $n \geq 6$, more than one inequality will be needed.

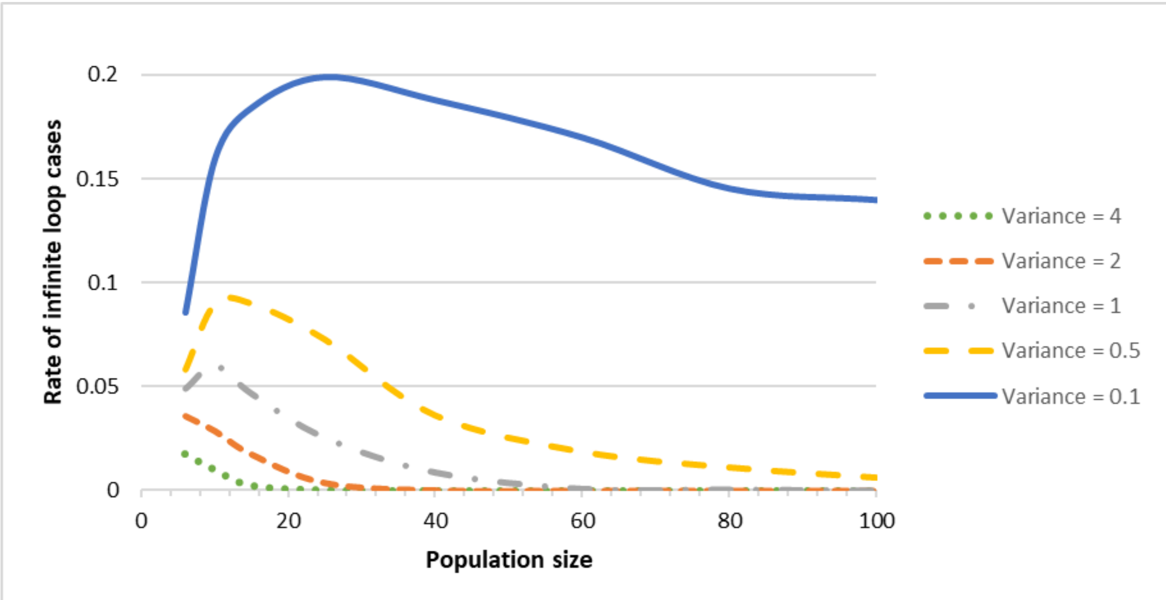
Given this difficulty, we investigated the case of $n \geq 6$ by means of simulations. We proceeded as follows. First, for a given population size taking one of eight particular values $n \in \{6, 10, 15, 25, 40, 60, 80, 100\}$, chosen so as to allow analysis of population sizes ranging from quite small to fairly large, we drew the incomes of the members of the population from a normal distribution characterized by a mean equal to 10 and a standard deviation equal to 3. We then allowed the individuals to move between region A and region B (incorporating the relative deprivation measure defined in section 3.1.1) until either a steady state distribution was reached, or a loop was encountered. This procedure was repeated 10,000 times with different incomes drawn from the same normal distribution. In Figure 1 we present the rate at which infinite loop cases were encountered during the 10,000 simulations for each chosen value of the eight population sizes. Although for small populations the rate of infinite loops was relatively low, an increase of population size yielded an increase in the rate of occurrence of infinite loops. For large populations this rate was close to 0.5. We revisited this observation later on (consult Claim 5). Similar results were obtained for other non-skewed distributions such as, for example, a uniform distribution, but for the sake of brevity we do not report them here.

FIGURE 1: The rate of occurrence of infinite loops: A normal distribution of incomes



A review of the case of a normal distribution of incomes leaves the impression that the dynamics of the movements between the two regions is more complicated than what appears to be in the case of the model presented in Section 3.1, in that a significant share of the initial income distributions does not lead to steady state at all. However, a normal distribution of incomes is not typical for real-world populations: real-world income distributions typically have positive skewness with long tails (Neal and Rosen, 2000). Intriguingly, for such distributions we obtain simple population dynamics: when the population is large, a steady state is typically reached after a single period (just as in the model of Section 3.1). To this end, we conducted simulations similar to the one for the normal distribution, employing instead the gamma distribution. We have chosen the gamma distribution partly because it has been frequently used for parametric analysis of income data (consult, for example, Salem and Mount, 1974), and partly because doing so enables us to ascertain simply and clearly what drives our results.[3] We considered again cases with $n \in \{6, 10, 15, 25, 40, 60, 80, 100\}$, and with parameters of the gamma distribution such that the expected value was always equal to 2, and the variance took the values of 4, 2, 1, 0.5, and 0.1. Keeping the expected value constant made it possible for us to investigate the pure effect of increased (or decreased) dispersion of incomes on the pattern of location choices. We obtained several illuminating results. In Figure 2 we present the rate at which infinite loop cases were encountered during the 10,000 simulations for each chosen value of population size and for each distribution variance. We see, first, that although initially an increase in the size of the population leads to an increase in the number of loop cases, after some threshold of population size is reached, the incidence of loop cases starts to decline, and it converges to zero. Second, in all cases loops are more likely to occur when the variance of the income distribution from which incomes are drawn is lower.

FIGURE 2: The rate of occurrence of infinite loop cases: A gamma distribution of incomes



That infinite loops occur more frequently when the incomes of a population are drawn from a gamma distribution with a lower variance suggests that it might be the case that loops are more likely in populations characterized by higher income equality. To further investigate this possibility, we compared the degree of income inequality in cases of steady states with the degree of income inequality in cases of infinite loops. To this end, we used the Gini index.[4] Specifically, we calculated the Gini index for each of the 10,000 simulations, and for all the analyzed population sizes and distribution variances. Then, the values of the index were averaged over the simulations in which a steady state distribution was achieved, and separately over the simulations in which an infinite loop was encountered. The results are summed up in Table 1.

TABLE 1: Averaged Gini indices for loop cases, and for steady state cases

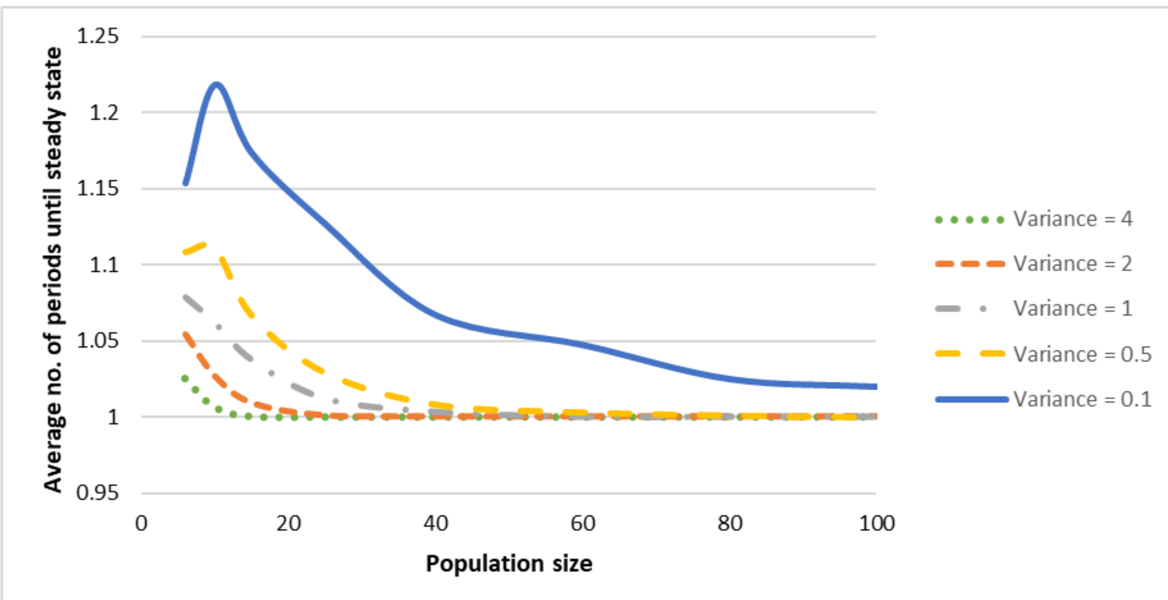
Variance = 4	6	10	15	25	40	60	80	100
Steady state	0.4206	0.4510	0.4682	0.4805	0.4877	0.4917	0.4935	0.4954
Loop	0.2635	0.2932	0.2981	0.2950	-	-	-	-
Variance = 2	**6**	**10**	**15**	**25**	**40**	**60**	**80**	**100**
Steady state	0.3152	0.3404	0.3506	0.3606	0.3661	0.3689	0.3700	0.3718
Loop	0.2243	0.2448	0.2592	0.2738	0.2718	-	-	-
Variance = 1	**6**	**10**	**15**	**25**	**40**	**60**	**80**	**100**
Steady state	0.2303	0.2499	0.2573	0.2636	0.2669	0.2686	0.2700	0.2708
Loop	0.1798	0.1966	0.2077	0.2203	0.2278	0.2374	0.2335	-
Variance = 0.5	**6**	**10**	**15**	**25**	**40**	**60**	**80**	**100**
Steady state	0.1650	0.1789	0.1856	0.1900	0.1922	0.1933	0.1943	0.1946
Loop	0.1385	0.1536	0.1618	0.1682	0.1753	0.1791	0.1787	0.1794
Variance = 0.1	**6**	**10**	**15**	**25**	**40**	**60**	**80**	**100**
Steady state	0.0742	0.0808	0.0837	0.0862	0.0872	0.0877	0.0880	0.0883
Loop	0.0681	0.0747	0.0795	0.0826	0.0846	0.0857	0.0866	0.0866

Note: For some infinite loop cases, no average Gini indices are displayed. This is so because for large populations, loops were not encountered at all (consult Figure 2).

From Table 1 we can infer that, indeed, on average cases in which infinite loops occur are characterized by smaller Gini indices than cases in which a steady state is reached. This is particularly visible for cases of a higher variance of the gamma distribution.

Looking at the average number of periods it takes to reach a steady state distribution (obviously in the cases in which a steady state is indeed reached) presented in Figure 3, we ask how much this number diverges from the result obtained in Section 3.1.1, namely convergence to a spatial steady state distribution in just one period. As can be seen in Figure 3, the lower the variance of the distribution, the higher the average number of periods it takes to reach a steady state, yet in all cases this number is not much greater than 1. Moreover, with a large population size, the average number of periods converges to 1. This finding is particularly revealing because in conjunction with the conclusions derived from Figure 2, it suggests that the results obtained from the highly simplified model of Section 3.1.1, namely that there always exists a steady state distribution of incomes which is achieved after one period of movement, are not too far removed from what happens in more complicated or more elaborate cases.

FIGURE 3: The average number of periods required to reach a steady state distribution: A gamma distribution of incomes



In addition, we have simulated distributions with properties that are similar to those of the gamma distribution, such as log-normal and inverse-Gaussian. These simulations yielded results that are nearly identical to the ones delivered by simulating the gamma distribution. We hasten to add that given Claim 5 below, this congruence is not surprising.

Why do the simulations yield the result that in large populations the number of periods it takes to reach a spatial steady state converges to 1, and no infinite loops occur?

Claim 5: Let X be a random variable with support over $[0,\infty)$, such that its expected value $E(X)$ exists and is finite, and such that $P(X > 2E(X)) > 0$. Let $(x_1, x_2, \ldots, x_n)$ be an ordered vector of incomes drawn from a probability distribution characterizing X. If we denote the outcome of reaching spatial steady state in one period by C, then $P(C)$ tends to 1 when n tends to infinity: namely for any $\varepsilon > 0$ there exists $n_0 > 0$ such that if $n > n_0$, then $P(C) > 1-\varepsilon$.

Proof: Consider a population of n individuals with different incomes. As always, in the first period of moving we will have the following distribution of the individuals between the two regions:

A	B
n	
	$n-1$
	$\vdots$
	1

Movement between the regions will cease if the relative deprivation of individual 1 (as he is the most relatively deprived individual) when in region B is lower than his possible relative deprivation in region A. When the income of individual *i* is x_i, this condition can be written as

$$\frac{x_n - x_1}{2} > \frac{\sum_{i=1}^{n-2}\left(x_{n-i} - x_1\right)}{n-1} \quad .$$

The right hand side of this inequality can be rewritten as

$$\frac{\sum_{i=1}^{n-2}\left(x_{n-i} - x_1\right)}{n-1} = \frac{\sum_{i=0}^{n-1} x_{n-i} - x_1 - x_n}{n-1} - \frac{n-2}{n-1}x_1 = \frac{\sum_{i=1}^{n} x_i}{n}\frac{n}{n-1} - \frac{x_1 + x_n}{n-1} - \frac{n-2}{n-1}x_1 .$$

Inserting this result into the inequality above and conducting several transformations leads to the inequality

$$\frac{1}{2}(x_1 + x_n + \frac{x_n - x_1}{n} - \frac{x_1}{n}) > \frac{\sum_{i=1}^{n} x_i}{n} .$$

Therefore:

$$P(C) = P\left(\frac{1}{2}(x_1 + x_n + \frac{x_n - x_1}{n} - \frac{x_1}{n}) > \frac{\sum_{i=1}^{n} x_i}{n}\right).$$

Let $\varepsilon > 0$ be fixed. There exists $\alpha > 0$ such that $P(X > 2E(X) + \alpha) > 0$. Let $P(X > 2E(X) + \alpha) = \beta > 0$. Then, for any $n > \log_{1-\beta}\frac{\varepsilon}{3}$, $P(x_n \le 2E(X) + \alpha) = (1-\beta)^n < \frac{\varepsilon}{3}$. Consequently, as $x_1 \ge 0$, $P(\frac{1}{2}(x_1 + x_n) \le E(X) + \frac{\alpha}{2}) < \frac{\varepsilon}{3}$.

Simultaneously, following the weak law of large numbers, there exists n_1 such that for $n > n_1$ $P\left(\frac{\sum_{i=1}^{n} x_i}{n} \geq E(X) + \frac{\alpha}{4}\right) < \frac{\varepsilon}{3}$. Additionally, let $n_2 \in \mathbb{N}$ be where for $n > n_2$ $P(\frac{x_1}{n} \geq \frac{\alpha}{2}) < \frac{\varepsilon}{3}$ (such n_2 always exists because $E(X)$ is finite and $P(x_1 < E(X)) > 0$).

If $\frac{x_1}{n} < \frac{\alpha}{4}$, $\frac{\sum_{i=1}^{n} x_i}{n} < E(X) + \frac{\alpha}{4}$, and $\frac{1}{2}(x_1 + x_n) > E(X) + \frac{\alpha}{2}$, then:

$$\frac{1}{2}(x_1 + x_n + \frac{x_n - x_1}{n} - \frac{x_1}{n}) > E(X) + \frac{\alpha}{2} + 0 - \frac{1}{2} \cdot \frac{\alpha}{2} > E(X) + \frac{\alpha}{4} > \frac{\sum_{i=1}^{n} x_i}{n}.$$

Consequently, for $n > n_0 = \max\{n_1, n_2, \log_{1-\beta} \frac{\varepsilon}{3}\}$:

$$P(C) = P\left(\frac{1}{2}(x_1 + x_n + \frac{x_n - x_1}{n} - \frac{x_1}{n}) > \frac{\sum_{i=1}^{n} x_i}{n}\right)$$

$$\geq P\left(\frac{x_1}{n} < \frac{\alpha}{4} \wedge \frac{\sum_{i=1}^{n} x_i}{n} < E(X) + \frac{\alpha}{4} \wedge \frac{1}{2}(x_1 + x_n) > E(X) + \frac{\alpha}{2}\right)$$

$$\geq 1 - P\left(\frac{x_1}{n} \geq \frac{\alpha}{4} \vee \frac{\sum_{i=1}^{n} x_i}{n} \geq E(X) + \frac{\alpha}{4} \vee \frac{1}{2}(x_1 + x_n) \leq E(X) + \frac{\alpha}{2}\right)$$

$$\geq 1 - P\left(\frac{x_1}{n} \geq \frac{\alpha}{4}\right) - P\left(\frac{\sum_{i=1}^{n} x_i}{n} \geq E(X) + \frac{\alpha}{4}\right) - P\left(\frac{1}{2}(x_1 + x_n) \leq E(X) + \frac{\alpha}{2}\right)$$

$> 1 - 3 \cdot \frac{\varepsilon}{3} = 1 - \varepsilon$. Q.E.D.

If the conditions assumed in Claim 5 are satisfied, then for large populations, moving will almost always cease after the first period, and there will be no (infinite) loops. These conditions are clearly not satisfied for every possible distribution. Nevertheless, they are satisfied for the gamma distribution that we have chosen, as well as for other distributions that share similar properties and characterize typical distribution of incomes in real-world populations, namely distributions having a long right tail. In particular, the condition $P(X > 2E(X)) > 0$ is not unrealistic as it states that there is a positive share of individuals who are more than twice as rich as the average individual.

From this discussion we can conclude that in large populations, the number of periods it takes to reach a steady state converges to one, because then the highest income in the population will be large, which will render it appealing for the lowest income individual to stay in region B. In region B, the distance between the second highest income and the lowest income will also be substantial, but the relative deprivation of the lowest income individual will be mitigated by the presence of many other individuals whose incomes are closer to his, albeit higher. Small populations are more likely to be more equal (consult Table 1), because with distributions such as gamma, the probability that there are no individuals whose income is twice the average income decreases with the size of the population. Therefore, it is more likely that individual 1 will move again after the first period, which will trigger a complex pattern of moves, leading possibly to an infinite loop.

6. ALLOWING INCOMES TO CHANGE

Up to now we have not allowed for the possibility that incomes, as such, can change. This has enabled us to concentrate on analyzing the pure effect of relative deprivation on the choice of location. When incomes can change, incorporating in the individuals' utility functions concern for low relative income together with a preference for (absolute) income can yield results that, in the absence of a preference for a better stance in social comparisons, could be considered somewhat counterintuitive. Consider the following example. An individual whose income is y, where $y > 2$, is in region A, where the income of the only other individual is $3y$. The individual can, alternatively, move to region B where the income awaiting him there will be $y-2$, and where no one has income higher than $y-2$. (Similarly, we can assume that moving to region B entails a cost of two units of income.) As assumed throughout this paper, the region where an individual is located constitutes the individual's region of social comparison. The individual likes absolute income and dislikes relative deprivation (which, again, we measure by the aggregate of income excesses divided by the size of the population), and assigns to these two terms in his utility function the weights of α and $-(1-\alpha)$, respectively, where $\alpha \in (0,1)$.

Definition 5: $u(x, RD) = \alpha x - (1-\alpha) RD$.

In this setting x denotes the individual's income, and RD denotes his relative deprivation, as per Definition 1. Then, if $\alpha < \frac{y}{y+2}$, the individual will prefer to move to region B.

Claim 6: For $y \to \infty$ the individual will always prefer to move to region B.

Proof: Defining $\frac{y}{y+2} \equiv \alpha_0$, it follows that $\frac{d\alpha_0}{dy} > 0$: as incomes rise, the constraint on α ($\alpha < \alpha_0$) for the individual's preference to move to region B becomes weaker. Because $\alpha_0 \equiv \frac{y}{y+2} = \frac{1}{1+\frac{2}{y}}$ we have that $\lim_{y \to \infty} \alpha_0 = 1$, so it follows that when incomes are fairly high, the constraint is not binding anymore. Q.E.D.

This result is intuitive because the higher is *y*, the less meaningful the difference between *y* and $y-2$, so leaving region A for region B involves an increasingly smaller relative loss of income, along with a significant (complete) reduction in relative deprivation.

The result reported above is robust to an alternative measure of relative deprivation. Suppose that instead of measuring RD as the sum of the income excesses divided by the size of the population, that it is measured as $\max\{\bar{x} - x, 0\}$, namely as the distance from below the mean. When the incomes in region A are y and $3y$, the mean income in the region is 2*y*, and the income distance of the region A's individual whose income is *y* from this mean income is *y*. This is the same *RD* as the *RD* obtained when measured by the income excesses divided by the size of the population: $\frac{1}{2}(3y - y) = y$.

Finally in this section, and as an informative example, we show how the model presented in Section 3 can be adjusted when moving from region A to region B involves a cost, $c > 0$. We retain the assumption that the individuals base their location decisions on the observed current state, without simultaneously forming expectations how other individuals will behave, and we follow the utility specification of Definition 5. In the general case in which to begin with (meaning in period 0) the individuals are in region A, an individual will move to region B only if the condition $\alpha < \frac{RD}{RD+c}$ is satisfied. This condition follows from comparison of the utilities in the two regions

$$U(i)\big|_{i\in A} = \alpha i - \left(1-\alpha\right) RD < \alpha\left(i-c\right) = U(i)\big|_{i\in B}.$$

Thus, and aligned with intuition, from an inspection of the condition we infer that an individual is more likely to move if he assigns a relatively low weight to income, if his relative deprivation in region A is high, and if the cost of moving is relatively low. If the cost of moving is relatively high, then some individuals from the bottom of the income hierarchy will not be able to afford to move, their high relative deprivation notwithstanding. We consider an example in which there are six individuals, $\alpha = 0.25$, and $c = 1$. The general case of n individuals, $\alpha \in (0,1)$, and $c > 0$ happens to be too complex to yield analytical solutions, although later on we comment on how different parameters impact on the results.

The sequence of movements leading to the steady state is depicted below. In period 0, the six individuals are in region A. It is clear that individual 6 has nothing to gain from moving to region B. With regard to individual 5, we calculate and compare his utilities in the two regions:

$$U(5)\big|_{5\in A} = \frac{1}{4}\cdot 5 - \frac{3}{4}\cdot\frac{1}{6} = \frac{9}{8} > \frac{1}{4}\cdot 4 = 1 = U(5)\big|_{5\in B}.$$

We infer that because the utility of individual 5 when in region A is higher than his utility would be if he were to move to region B, he will stay in region A. However, individuals 1, 2, 3, and 4 will move to region B. The reason is that for each of them, the relative deprivation experienced in region A burdens more than bearing the cost of moving to region B where the consequent relative deprivation relief counts heavily.

After the first period, no individual who is in region B has an incentive to move back to region A: individual 4 has nothing to gain from moving back; individuals 3 and 2 will not only have to bear a cost of moving back, but will also have their relative deprivation increasing, so they will stay; and individual 1 cannot afford to move. On the other hand, now that individuals 1, 2, 3, and 4 left region A, individual 5 is more relatively deprived staying there. The condition $\alpha < \frac{RD}{RD+c}$ is satisfied for him and, therefore, he will elect to move. In period 2 we reach then the same distribution as in the main model of Section 3, with one individual staying in region A, and the rest of the individuals in region B. This distribution constitutes a steady state: with individual 6 in region A and the remainder of the population in region B, no individual in region B has an incentive to move back, because if any of them were to do so, the resulting relative deprivation will be higher. Thus, in this example the result of a steady state outcome with a sharp bifurcation continues to hold, although the number of periods it takes to reach the steady state is bigger than one. Here it is two periods: because of the positive cost of moving, relocating is less attractive to individual 5 who is initially not much relatively deprived. However, in the wake of the departure of individuals 1, 2, 3, and 4, the increased relative deprivation of individual 5 overrides the cost of moving, so he ends up moving too. We hasten to add that in the general case, the outcome can differ from the one reported here, especially so if the cost of movement is higher and / or if the individuals attach a higher weight to income in their utility function.

A	B
6	
5	
4	
3	
2	
1	

→

A	B
6	
5	
	4
	3
	2
	1

→

A	B
6	
	5
	4
	3
	2
	1

7. DISCUSSION

7.1 An alternative initial distribution

Until now we have assumed that to begin with all the individuals are in region A, and that region B is empty. It is tempting to inquire what happens if, instead, we assume that the initial distribution between the two regions of individuals $1, 2, \ldots, n$ with, correspondingly, incomes x_i, $i = 1, 2, ..., n$, such that $x_1 < x_2 < \ldots < x_n$, is arbitrary. In particular, does Claim 5 still hold? As it turns out, it does not. In the case of a larger set of income vectors $x = (x_1, x_2, \ldots, x_n)$, the dynamics is more complicated than in the case in which to begin with all the individuals are in region A, with the outcome depending on the initial distribution of the individuals, not only on their incomes.

From the proof of Claim 5 we know that movement between the regions will cease after just one period and will stabilize at a steady state in which all the individuals but *n* are in region B, and that this outcome obtains when to begin with all the individuals are in region A and $\frac{1}{2}(x_1 + x_n + \frac{x_n - x_1}{n} - \frac{x_1}{n}) > \frac{\sum_{i=1}^{n} x_i}{n} = \overline{x}$, in particular when $x_n > 2\overline{x}$, which is essential for Claim 5 to hold (the assumption $P(X > 2E(X)) > 0$ practically guarantees that for sufficiently large n, $x_n > 2\overline{x}$). Thus, we know that if, for example, $n = 4$, $x_1 = 1$, $x_2 = 2$, $x_3 = 15$, and $x_4 = 20$, and to begin with these four individuals are in region A, then $x_4 = 20 > 2 \cdot \frac{38}{4} = 2\overline{x}$, and the evolving dynamics is depicted by the sequence:

A	B
20	
15	
2	
1	

$\rightarrow$

A	B
20	
	15
	2
	1

However, this is not the only possibility. For example, when to begin with individual 3 ($x_3 = 15$) is in region B, the evolving dynamics results in a different steady state:

A	B
20	
	15
2	
1	

In addition, it is possible that the dynamics for the same vector of incomes will not converge to a steady state at all. If to begin with individuals 4 and 3 are in different regions and individuals 2 and 1 are in different regions, then the process repeats itself ad infinitum, and a steady state is not reached:

A	B
20	
	15
2	
	1

→

A	B
20	
	15
	2
1	

→

A	B
20	
	15
2	
	1

→

A	B
20	
	15
	2
1	

→ …

This example reveals that, when to begin with not all the individuals are in the same region, the condition $x_n > 2\bar{x}$ no longer guarantees convergence of the evolving dynamics, and even if convergence obtains, there can be multiple steady states. Thus, the potential sufficient condition for convergence to obtain will need to take into account not only the income vectors, but also the initial distribution of the individuals between the two regions.

When the income differences between all pairs of adjacent individuals are the same, sometimes it is not difficult to accommodate a constellation in which not all the individuals are initially in region A, but sometimes it can lead to different outcomes than the one presented in Claims 1 and 2. For example, suppose that there are two regions, A and B, and that to begin with individuals 5, 4, and 3 are in region A, and individuals 2 and 1 are in region B. Once movement between the regions is allowed, individuals 4 and 3 will move to region B, and the steady state distribution will be for individual 5 to be in region A, with the remainder of the individuals in region B. This outcome obtains because there was no incentive for individuals 2 and 1 to move, neither when individuals

4 and 3 were in region A, nor after individuals 4 and 3 moved to region B. However, if to begin with individuals 5, 2, and 1 are in region A, and individuals 4 and 3 are in region B, then no individual has an incentive to move, and this distribution in and by itself constitutes a steady state. This steady state differs from the steady state in which individual 5 is in region A, and the other individuals are in region B. Thus, Claim 1 also need not hold for arbitrary initial distributions of the individuals between the regions. These examples help informing us as follows. For $n \leq 4,$ we will always obtain a steady state as per Claims 1 and 2. For $n = 5,$ we will always obtain a steady state, but the steady state may not be as per Claims 1 and 2. For $n \geq 6,$ we can obtain a steady state as per Claims 1 and 2, or we may not obtain a steady state at all as when, for example, to begin with individuals 6, 4, 3, and 1 are in region A, and individuals 5 and 2 are in region B. An analysis for $n \leq 20$ of all possible initial distributions of the individuals between the two regions reveals that only for $n = 5$ there is one initial distribution which leads to a different steady state than the steady state obtained as per Claims 1 and 2. For $6 \leq n \leq 12$, there are initial distributions that do not lead to a steady state at all. For $13 \leq n \leq 20$, all the initial distributions lead to the same steady state as per Claims 1 and 2.

While it is beyond the scope of this paper, a more thorough analysis of the dynamics of movement for any initial distribution of the individuals between the two regions would be an intriguing topic for follow-up research.

7.2 More than two regions

Hitherto we have studied a setting in which there are two regions. What happens if there are more than two regions?

Consider again the case of n individuals. Let the income of individual i be x_i, $i=1,2,...,n$, and let $x_1 < x_2 < \ldots < x_n$. Let there be k regions: $A_1, A_2, \ldots, A_k$ and to begin with let all the individuals be in a region A_1. It is quite obvious that if there are at least as many (identical) regions as there are individuals, then each individual will be able to experience zero relative deprivation, occupying his own region all by himself. Therefore, we assume next that $2 < k < n$.

To understand the dynamics of movement between k regions, we need to establish how individuals choose to move to one of a number of equally attractive regions. During the first period of movement we only know that individual n stays in A_1, and that the other individuals move, but we do not know where to because from their point of view, regions $A_2, A_3, \ldots, A_n$ are identical.

If the individuals choose randomly between moving to equally attractive regions, then we cannot determine the outcome of such dynamics. As an example, we consider the case of three regions: *A*, *B*, and *C*, and of $n=5$ individuals such that $x_1 = 1$, $x_2 = 2$, $x_3 = 15$, $x_4 = 20$, and $x_5 = 100$. In the first period, individual 5 stays in region A, and individuals 1, 2, 3, and 4 randomly move to region B or to region C. In the subsequent periods, individual 5 always remains in region A, no other individual ever goes back to region A (because the relative deprivation arising from a comparison with 5 is too high), and the dynamics of movement of individuals 1 to 4 between regions B and C is exactly the same as the one considered in Subsection 7.1 where the random allocation in period 1 generates the initial condition. In particular, if in the first period individuals 1, 2, 3, and 4 move to region B, then the dynamics converges to the steady state:

A	B	C
100		
	20	
		15
		2
		1

and if in the first period individuals 1, 2, and 4 move to region B while individual 3 moves to region C, then the system stays in the steady state:

A	B	C
100		
	20	
		15
	2	
	1	

Finally, if in the first period individuals 2 and 4 move to region B and individuals 1 and 3 move to region C, then a steady state is never reached.

Thus, a deterministic rule is needed in order to decide between equally attractive destinations. However, there are many possible rules, and each of them leads to dynamics which yields an outcome that is at least as complex as the one presented in Section 5. Therefore, we consider only a simple generalization of the dynamics of Section 3.

Assume that for every i, $x_i = i$, and that when choosing between two equally attractive regions A_l and A_m the individuals choose the one with the lowest index, namely A_l if $l < m$ and A_m otherwise. This assumption is reasonable when regions with lower indices are more accessible than regions with higher indices, for example when for

l, m such that $1 < l < m$ an individual needs to go through region A_l when moving from A_1 to A_m. Then, the dynamics is analogous to the one portrayed in Subsection 3.1.1: in the first period individual n stays in region A_1 and the other individuals move to region A_2. In the second period, individual n stays in region A_1, individual $n-1$ stays in region A_2, and the other of individuals move to region A_3. Repeating the procedure, after $k-1$ periods, the dynamics converges to a steady state where for $j < k$ only individual $n+1-j$ stays in region A_j, and the other individuals stay in region A_k. The reasoning behind this process is analogous to that of the proof of Claim 1. For example, when $n = 5$ and $k = 3$, then after two periods the following steady state is reached:

A_1	A_2	A_3
5		
	4	
		3
		2
		1

8. CONCLUSION

We have shown how dissatisfaction arising from having low relative income can influence the choice of location. The ensuing dynamics can take a variety of forms. For any *n* when incomes are equally spaced, a steady state spatial distribution will be reached under alternative indices of cardinally measured relative deprivation, with the end distribution being the same, even though the dynamic paths leading to the end distribution differ. When incomes are not equally spaced, we formulated conditions such that, under relative deprivation, a steady state spatial distribution will be reached for any

$n \leq 5.$ Given the difficulty in analyzing directly cases of $n \geq 6,$ we resorted to simulations. We obtained several results: first, the probability of reaching a steady state spatial distribution approaches 1 as the size of the population increases. Second, the average number of periods it takes to reach a steady state also converges to 1. Third, the incidence of cyclical moves (with a steady state not reached) is more likely in populations with lower income inequality, as measured by both the Gini index and by the variance of the distribution of incomes.[5] This result suggests a particular testable hypothesis: a more stable pattern of movement between regions exists when the interregional income variation is larger.

ACKNOWLEDGMENTS

We are indebted to a referee for kind words and illuminating suggestions, and to Steven Brakman for advice and guidance. Grzegorz Kosiorowski's research was partially financed from the funds granted to the Faculty of Finance and Law at Cracow University of Economics within the framework of the subsidy for the maintenance of the Faculty's research potential.

FOOTNOTES

[1] Hitherto, in order to simplify the notation, we did not represent the relative deprivation function as a function of the incomes of all the individuals.

[2] By "analogous" we mean that the equivalent of Lemmas 1, 2, and 3, as well as the analysis for any natural number *n* for the case of relative deprivation defined as the income distance from below the mean income, are also of the form: "For any natural

number $n > 3$, if some set of inequalities is satisfied then the distribution of the individuals between the two regions will perpetually change. Otherwise, a spatial steady state will be reached." The proofs of the equivalent lemmas are also similar. Naturally, the number and the exact form of the inequalities may differ between the two definitions of relative deprivation.

[3] More sophisticated distributions were proposed, inter alia, by McDonald (1984), and more recently by Chatterjee et al. (2007).

[4] Mathematically, the Gini index is equivalent to total relative deprivation divided by total income. This forges a link with the measure of the distaste for low relative income.

[5] This implication is valid for distributions similar to the gamma distribution, namely for distributions that have positive skewness with long tail. It does not hold for other distributions such as the uniform distribution.

APPENDIX A: PROOFS

Lemma 1. When $n=3$, a spatial steady state with individual 3 in region A, and individuals 1 and 2 in region B will always be reached.

Proof.

1st period: Individual 1 (whose income is x_1) and individual 2 (whose income is x_2) move to region B so as to shed off their relative deprivation, while individual 3 (whose income is x_3) remains in region A:

A	B
3	
	2
	1

2nd period: Because individuals 2 and 3 are not relatively deprived, they will not have an incentive to move, and they are thus of no further "interest" to us. The only individual to consider then is individual 1. For this individual when he is in region B:

$$RD(1)\big|_{1\in B}=\frac{1}{2}(x_2-x_1)\,.$$

If individual 1 were to return to region A:

$$RD(1)\big|_{1\in A}=\frac{1}{2}(x_3-x_1)\,.$$

Because $x_2<x_3$, individual 1 will prefer to stay in region B. Thus, after the first period no individual has an incentive to move, so the distribution reached in the first period is the spatial steady state distribution. Q.E.D.

Lemma 2. When $n=4$, the distribution of the individuals between the two regions will perpetually change and a spatial steady state will not be reached iff $2x_4+x_2<3x_3$ and $3x_4+x_1<2(x_3+x_2)$.

Proof.

1st period: Individuals 1, 2, and 3 move to region B to get rid of their relative deprivation, and individual 4 remains in region A:

A	B
4	
	3
	2
	1

2nd period: Individuals 3 and 4 are not relatively deprived, and therefore they will not relocate. For individual 1 when he is in region B:

$$RD(1)\big|_{1\in B}=\frac{1}{3}(x_3+x_2-2x_1)\,.$$

If individual 1 were to return to region A:

$$RD(1)\big|_{1\in A}=\frac{1}{2}(x_4-x_1)\,.$$

We have that $RD(1)\big|_{1\in A}\geq RD(1)\big|_{1\in B}\Leftrightarrow 3x_4+x_1\geq 2(x_3+x_2)$. Therefore, if the inequality $3x_4+x_1\geq 2(x_3+x_2)$ is satisfied, individual 1 will prefer to stay in region B.

For individual 2 when he is in region B:

$$RD(2)\big|_{2\in B}=\frac{1}{3}(x_3-x_2)\,.$$

If individual 2 were to return to region A:

$$RD(2)\big|_{2\in A} = \frac{1}{2}(x_4 - x_2).$$

Because $x_3 < x_4$, individual 2 will prefer to stay in region B.

Therefore, if $3x_4 + x_1 \geq 2(x_3 + x_2)$ the spatial distribution reached after the first period is a steady state distribution, as no individual has an incentive to move. However, if the condition $3x_4 + x_1 \geq 2(x_3 + x_2)$ does not hold, then the dynamics of locational choices is not brought to a halt because individual 1 has an incentive to move to region A. Thus, assuming that $3x_4 + x_1 < 2(x_3 + x_2)$, the distribution of the individuals between the two regions in the second period is

A	B
4 1	 3 2

3rd period: From now on, we assume that $3x_4 + x_1 < 2(x_3 + x_2)$ because when the individuals' incomes do not observe this inequality, period three will not occur at all, and the location dynamics will come to a halt after the first period.

Individuals 3 and 4 are not relatively deprived, so they will not move. From the calculations in the preceding step we know that individual 1 prefers his current location in region A over returning to region B. Then the only possible mover in this period is individual 2. If individual 2 remains in region B:

$$RD(2)\big|_{2\in B} = \frac{1}{2}(x_3 - x_2).$$

If individual 2 were to return to region A:

$$RD(2)\big|_{2\in A} = \frac{1}{3}(x_4 - x_2)\,.$$

We have that $RD(2)\big|_{2\in A} \geq RD(2)\big|_{2\in B} \Leftrightarrow 2x_4 + x_2 \geq 3x_3$. Therefore, if the inequality $2x_4 + x_2 \geq 3x_3$ is satisfied, individual 2 will prefer to stay in his current location, namely in region B. Then, the distribution reached in the second period (with individuals 1 and 4 in region A, and individuals 2 and 3 in region B) constitutes a steady state.

However, if $2x_4 + x_2 < 3x_3$, individual 2 will prefer to move to region A, and the individuals will be distributed between the two regions as follows:

A	B
4	
	3
2	
1	

4th period: From now on, we assume that $2x_4 + x_2 < 3x_3$ and that $3x_4 + x_1 < 2(x_3 + x_2)$. Otherwise, the dynamics of moves between locations will come to a halt earlier, and the fourth period will not occur at all. Analogically to the preceding steps, we can infer that the only possible mover in this period is individual 1. If individual 1 remains in region A:

$$RD(1)\big|_{1\in A} = \frac{1}{3}(x_4 + x_2 - 2x_1)\,.$$

If individual 1 were to return to region B:

$$RD(1)\big|_{1\in B} = \frac{1}{2}(x_3 - x_1)\,.$$

We have that $3x_4 + x_1 < 2(x_3 + x_2)$, thus $-2x_4 + x_1 < -3x_3 + 2x_2$ (because $x_4 > x_3$) which can be rearranged as $3x_3 - 3x_1 < 2x_4 + 2x_2 - 4x_1$, and finally as $\frac{1}{2}(x_3 - x_1) < \frac{1}{3}(x_4 + x_2 - 2x_1)$. This last inequality implies that $RD(1)\big|_{1\in B} < RD(1)\big|_{1\in A}$. Therefore, individual 1 moves again to region B, and the distribution of the individuals between the two regions is:

A	B
4	
	3
2	
	1

5th period: Analogically to the preceding steps, we can infer that the only possible mover in this period is individual 2. If individual 2 remains in region A:

$$RD(2)\big|_{2\in A} = \frac{1}{2}(x_4 - x_2).$$

If individual 2 were to return to region B:

$$RD(2)\big|_{2\in B} = \frac{1}{3}(x_3 - x_2).$$

Because $x_3 < x_4$, individual 2 will prefer to move to region B. Now, the process reverts back to the configuration that prevailed after the relocation moves in the first period. If $2x_4 + x_2 < 3x_3$ and $3x_4 + x_1 < 2(x_3 + x_2)$, this loop repeats itself ad infinitum. As we concluded at the end of the analysis of the 2nd and 3rd periods, if any of these inequalities is not satisfied, the system reaches a spatial steady state. Q.E.D.

Lemma 3. When $n = 5$, the distribution of the individuals between the two regions will perpetually change and a spatial steady state will not obtain iff $2x_5 + x_1 < x_4 + x_3 + x_2$ and $x_5 + x_2 < x_4 + x_3$ and $x_5 + x_3 < 2x_4$.

Proof.

1st period: Individuals 1, 2, 3, and 4 move to region B to get rid of their relative deprivation, and individual 5 remains in region A.

A	B
5	
	4
	3
	2
	1

2nd period: Individuals 5 and 4 are not relatively deprived. Therefore, they will not move. For individual 3 when he is in region B:

$$RD(3)\big|_{3\in B} = \frac{1}{4}(x_4 - x_3)\ .$$

If individual 3 were to return to region A:

$$RD(3)\big|_{3\in A} = \frac{1}{2}(x_5 - x_3)\,.$$

Because $x_4 < x_5$, individual 3 will prefer to stay in region B.

For individual 2 when he is in region B:

$$RD(2)\big|_{2\in B} = \frac{1}{4}(x_4 + x_3 - 2x_2)\,.$$

If individual 2 were to return to region A:

$$RD(2)\big|_{2\in A} = \frac{1}{2}(x_5 - x_2).$$

Because $x_3 + x_4 < 2x_5$, we have that $x_3 + x_4 - 2x_2 < 2x_5 - 2x_2$ or that $\frac{1}{4}(x_4 + x_3 - 2x_2) < \frac{1}{2}(x_5 - x_2)$. Therefore, individual 2 will prefer to stay in region B.

For individual 1 when he is in region B:

$$RD(1)\big|_{1\in B} = \frac{1}{4}(x_4 + x_3 + x_2 - 3x_1).$$

If individual 1 were to return to region A:

$$RD(1)\big|_{1\in A} = \frac{1}{2}(x_5 - x_1).$$

We have that $RD(1)\big|_{1\in A} \geq RD(1)\big|_{1\in B} \Leftrightarrow 2x_5 + x_1 \geq x_4 + x_3 + x_2$. Therefore, if the inequality $2x_5 + x_1 \geq x_4 + x_3 + x_2$ is satisfied, individual 1 will prefer to stay in region B, and then the distribution reached after period one is a steady state distribution because no individual has an incentive to move.

However, when $2x_5 + x_1 < x_4 + x_3 + x_2$, individual 1 moves to region A, and the distribution of the individuals between the two regions is:

A	B
5	
	4
	3
	2
1	

3rd period: From now on, we assume that $2x_5 + x_1 < x_4 + x_3 + x_2$. Individuals 5 and 4 will not move. From the calculations in the preceding step we know that individual 1 prefers his current location in region A over returning to region B. The only possible movers in this period are individuals 2 and 3. If individual 3 remains in region B:

$$RD(3)\big|_{3\in B} = \frac{1}{3}(x_4 - x_3)\,.$$

If individual 3 were to return to region A:

$$RD(3)\big|_{3\in A} = \frac{1}{3}(x_5 - x_3)\,.$$

Because $x_4 < x_5$, individual 3 will prefer to stay in region B. If individual 2 remains in region B:

$$RD(2)\big|_{2\in B} = \frac{1}{3}(x_4 + x_3 - 2x_2)\,.$$

If individual 2 were to return to region A:

$$RD(2)\big|_{2\in A} = \frac{1}{3}(x_5 - x_2)\,.$$

We have that $RD(2)\big|_{2\in B} \le RD(2)\big|_{2\in A} \Leftrightarrow x_4 + x_3 \le x_5 + x_2$. Therefore, if the inequality $x_4 + x_3 \le x_5 + x_2$ is satisfied, individual 2 will prefer to stay in region B. Then, the distribution reached in the second period (with individuals 1 and 5 in region A, and individuals 2, 3, and 4 in region B) constitutes a steady state distribution.

However, if $x_4 + x_3 > x_5 + x_2$, individual 2 moves to region A, and we obtain the following distribution of the individuals between the two regions:

A	B
5	
	4
	3
2	
1	

4th period: From now on, we assume that $2x_5 + x_1 < x_4 + x_3 + x_2$ and that $x_4 + x_3 > x_5 + x_2$. Individuals 5 and 4 will not move. From the calculations in the preceding step we know that individual 2 prefers his current location in region A over returning to region B. The only possible movers in this period are individuals 1 and 3. If individual 1 remains in region A:

$$RD(1)\big|_{1\in A} = \frac{1}{3}(x_5 + x_2 - 2x_1)\ .$$

If individual 1 were to move to region B:

$$RD(1)\big|_{1\in B} = \frac{1}{3}(x_4 + x_3 - 2x_1)\ .$$

Because $x_4 + x_3 > x_5 + x_2$, individual 1 will prefer to stay in region A.

If individual 3 remains in region B:

$$RD(3)\big|_{3\in B} = \frac{1}{2}(x_4 - x_3)\ .$$

If individual 3 were to return to region A:

$$RD(3)\big|_{3\in A} = \frac{1}{4}(x_5 - x_3)\ .$$

We have that $RD(3)\big|_{3\in A} \geq RD(3)\big|_{3\in B} \Leftrightarrow x_5 + x_3 \geq 2x_4$. Therefore, if the inequality $x_5 + x_3 \geq 2x_4$ is satisfied, individual 3 will prefer to stay in region B. Then, the distribution reached in period 3 (with individuals 1, 2, and 5 in region A, and individuals 3 and 4 in region B) constitutes a steady state distribution.

However, when the opposite holds, namely when $x_5 + x_3 < 2x_4$, individual 3 moves to region A. In this case, the distribution of the individuals between the two regions is:

A	B
5	
	4
3	
2	
1	

5th period: From now on, we assume that $2x_5 + x_1 < x_4 + x_3 + x_2$, that $x_4 + x_3 < x_5 + x_2$, and that $x_5 + x_3 < 2x_4$. Individuals 5 and 4 will not move. From the calculations in the preceding step we know that individual 3 prefers his current location in region A over returning to region B. The only possible movers in this period are individuals 1 and 2.

If individual 2 remains in region A:

$$RD(2)\big|_{2\in A} = \frac{1}{4}(x_5 + x_3 - 2x_2).$$

If individual 2 were to move to region B:

$$RD(2)\big|_{2\in B} = \frac{1}{2}(x_4 - x_2).$$

We have that $x_5 + x_3 < 2x_4$, thus $x_5 + x_3 - 2x_2 < 2x_4 - 2x_2$ and finally $\frac{1}{4}(x_5 + x_3 - 2x_2) < \frac{1}{2}(x_4 - x_2)$. This last inequality implies that $RD(2)|_{2\in A} < RD(2)|_{2\in B}$, so individual 2 will prefer to stay in region A.

If individual 1 remains in region A:

$$RD(1)|_{1\in A} = \frac{1}{4}(x_5 + x_3 + x_2 - 3x_1).$$

If individual 1 were to move to region B:

$$RD(1)|_{1\in B} = \frac{1}{2}(x_4 - x_1).$$

We have that $2x_5 + x_1 < x_4 + x_3 + x_2$ and $x_5 > x_4$, thus $2x_4 + x_1 < x_5 + x_3 + x_2$. After subtracting $3x_1$ from both sides of the last inequality, we get that $2x_4 - 2x_1 < x_5 + x_3 + x_2 - 3x_1$, and finally $\frac{1}{2}(x_4 - x_1) < \frac{1}{4}(x_5 + x_3 + x_2 - 3x_1)$. This last condition implies that $RD(1)|_{1\in B} < RD(1)|_{1\in A}$, so individual 1 will prefer to move to region B. Consequently, the distribution of the individuals between the regions becomes:

A	B
5	
	4
3	
2	
	1

6th period: Individuals 5 and 4 will not move. From the calculations in the preceding step we know that individual 1 prefers his current location in region B over returning to region A. The only possible movers in this period are individuals 3 and 2.

If individual 3 remains in region A:

$$RD(3)\big|_{3\in A} = \frac{1}{3}(x_5 - x_3)\,.$$

If individual 3 were to move to region B:

$$RD(3)\big|_{3\in B} = \frac{1}{3}(x_4 - x_3)\,.$$

Because $x_4 < x_5$, individual 3 will prefer to move to region B.

If individual 2 remains in region A:

$$RD(2)\big|_{2\in A} = \frac{1}{3}(x_5 + x_3 - 2x_2)\,.$$

If individual 2 were to move to region B:

$$RD(2)\big|_{2\in B} = \frac{1}{3}(x_4 - x_2)\,.$$

We have that $x_4 < x_5$ and $x_2 < x_3$, thus $x_4 + x_2 < x_5 + x_3$. Upon subtracting $2x_2$ from both sides of the last inequality, we get that $x_4 - x_2 < x_5 + x_3 - 2x_2$, implying that $RD(2)\big|_{2\in B} < RD(2)\big|_{2\in A}$, so individual 2 will prefer to move to region B.

With both individuals 3 and 2 moving to region B, the distribution of the individuals between the two regions in this period is given by:

A	B
5	
	4
	3
	2
	1

This distribution replicates the distribution reached in period 1. And the loop repeats itself ad infinitum, assuming that the individuals' incomes are such that the three following conditions hold: $2x_5 + x_1 < x_4 + x_3 + x_2$, $x_4 + x_3 > x_5 + x_2$, and $x_5 + x_3 < 2x_4$. If any of these conditions is not satisfied, the spatial distribution of the individuals between the two regions will reach a steady state. Q.E.D.

APPENDIX B: CONSTRUCTION OF THE MEASURE OF RELATIVE DEPRIVATION PRESENTED IN DEFINITION 1

For the purpose of constructing a measure, a natural starting point is the work of Runciman (1966), who argued that an individual has an unpleasant sense of being relatively deprived when he lacks a desired good and perceives that others with whom he naturally compares himself possess that good. Runciman (1966, p. 19) writes as follows: "The more people a man sees promoted when he is not promoted himself, the more people he may compare himself with in a situation where the comparison will make him feel deprived," implying that the deprivation from not having, say, income y is an increasing function of the fraction of people in the individual's reference group who have y. To aid intuition and for the sake of concreteness, we resort to income-based comparisons, meaning that an individual feels relatively deprived when others in his reference group earn more than he does. It is assumed implicitly here that the earnings of others are publicly known. Alternatively, we can think of consumption, which might be more publicly visible than income, although these two variables can reasonably be assumed to be closely related.

As an illustration of the relationship between the fraction of people possessing income y and the deprivation of an individual lacking y, consider a population (reference group) of six individuals with incomes $\{1,2,6,6,6,8\}$. Imagine a furniture store that in three distinct departments sells chairs, armchairs, and sofas. An income of 2 allows you to buy a chair. To be able to buy an armchair, you need an income that is a little bit higher than 2. To buy any sofa, you need an income that is a little bit higher than 6. Thus, when you go to the store and your income is 2, what are you "deprived of?" The answer is “of armchairs” and “of sofas.” Mathematically, this deprivation can be represented by $P(Y>2)(6-2)+P(Y>6)(8-6)$, where $P(Y>y_i)$ stands for the fraction of those in the population whose income is higher than y_i, for $y_i = 2,6$. The reason for this representation is that when you have an income of 2, you cannot afford anything in the department that sells armchairs, and you cannot afford anything in the department that sells sofas. Because not all those who are to your right in the ascendingly ordered income distribution can afford to buy a sofa, but they can all afford to buy armchairs, a breakdown into the two (weighted) terms $P(Y>2)(6-2)$ and $P(Y>6)(8-6)$ is needed. This way, we get to the very essence of the measure of *RD* presented in this paper: we take into account the fraction of the reference group (population) who possess some good which you do not, and we weigh this fraction by the “excess value” of that good. Because income enables an individual to afford the consumption of certain goods, we refer to comparisons based on income.

Formally, let $y=(y_1,...,y_m)$ be the vector of incomes in population *N* of size *n* with relative incidences $p(y)=\left(p(y_1),...,p(y_m)\right)$, where $m \le n$ is the number of distinct income levels in *y*, where *n* and *m* are natural numbers. The *RD* of an individual earning y_i is defined as the weighted sum of the excesses of incomes higher than y_i such that each excess is weighted by its relative incidence, namely

$$RD_N(y_i) \equiv \sum_{y_k > y_i} p(y_k)(y_k - y_i)\,. \tag{B1}$$

In the example given above with income distribution {1,2,6,6,6,8}, we have that the vector of incomes is $y=(1,2,6,8)$, and that the corresponding relative incidences are $p(y)=(1/6,1/6,3/6,1/6)$. Therefore, the *RD* of the individual earning 2 is

$\sum_{y_k > y_i} p(y_k)(y_k - y_i) = p(6)(6-2) + p(8)(8-2) = \frac{3}{6} \cdot 4 + \frac{1}{6} \cdot 6 = 3$. By similar calculations, we have that the *RD* of the individual earning 1 is higher at $3\frac{5}{6}$, and that the *RD* of each of the individuals earning 6 is lower at $\frac{1}{3}$.

We expand the vector y to include incomes with their possible respective repetitions, that is, we include each y_i as many times as its incidence dictates, and we assume that the incomes are ordered, that is, $y = (y_1, ..., y_n)$ such that $y_1 \le y_2 \le ... \le y_n$. In this case, the relative incidence of each y_i, $p(y_i)$, is $1/n$, and (B1), defined for $i = 1,...,n-1$, becomes

$$RD_N(y_i) \equiv \frac{1}{n} \sum_{k=i+1}^{n} (y_k - y_i). \tag{B2}$$

This (B2) expression is the basis of Definition 1.

REFERENCES

Basarir, H. (2012). Is it all relative? The role of wealth in the migration decision: Evidence from Indonesia (Working Paper). University of York.

Chatterjee, A., Yarlagadda, S. and Chakrabarti, B. K. (2007). *Econophysics of wealth distributions: Econophys-Kolkata I*. New York: Springer Science & Business Media.

Czaika, M. (2012). Internal versus international migration and the role of multiple deprivation: Some evidence from India. *Asian Population Studies* 8(2), 125-149.

Flippen, C. (2013). Relative deprivation and internal migration in the United States: A comparison of black and white men. *American Journal of Sociology*, 118(5), 1161-1198.

Frey, B. S. and Stutzer, A. (2002). *Happiness and economics: How the economy and institutions affect human well-being*. Princeton and Oxford: Princeton University Press.

Hyll, W. and Schneider, L. (2014). Relative deprivation and migration preferences. *Economics Letters*, 122(2), 334-337.

Jagger, P., Shively, G. and Arinaitwe, A. (2012). Circular migration, small-scale logging, and household livelihoods in Uganda. *Population and Environment*, 34(2), 235-256.

Kafle K., Benfica, R., and Winters, P. (2018). Does relative deprivation induce migration? International Fund for Agricultural Development Research Series 21, Rome.

McDonald, J. B. (1984). Some generalized functions for the size distribution of income. *Econometrica*, 52(3), 647-663.

Neal, D. and Rosen, S. (2000). Theories of the distribution of earnings. In A.B. Atkinson, F. Bourguigon (Eds.), *Handbook of Income Distribution* (1). Amsterdam: North Holland, 379-427.

Quinn, M. A. (2006). Relative deprivation, wage differentials and Mexican migration. *Review of Development Economics*, 10(1), 135-153.

Runciman, W. G. (1966). *Relative Deprivation and Social Justice*. Berkeley: University of California Press.

Salem, A. B. Z. and Mount, T. D. (1974). A convenient descriptive model of income distribution: The gamma density. *Econometrica*, 42(6), 1115-1127.

Stark, O. (2013). Stressful integration. *European Economic Review*, 63, 1-9.

Stark, O., Micevska, M. and Mycielski, J. (2009). Relative poverty as a determinant of migration: Evidence from Poland. *Economics Letters*, 103(3), 119-122.

Stark, O. and Taylor, E. J. (1991). Migration incentives, migration types: The role of relative deprivation. *Economic Journal*, 101, 1163-1178.

Vernazza, D. (2013). Does absolute or relative income motivate migration? London School of Economics, Mimeo.

Walker, I. and Smith, H. J. (2002). *Relative deprivation: Specification, development, and integration*. Cambridge: Cambridge University Press.