WIDER Working Paper 2020/12

Social mobility and inequality between groups

Patricia Funjika¹,² and Rachel M. Gisselquist³

February 2020
Abstract: The relationship between social mobility and inequality is well studied in the literature, but far greater attention has been paid to ‘vertical’ than to ‘horizontal’ inequality. This paper focuses on mobility and horizontal inequality between ethnic, racial, and culturally-defined groups. Not only is persistence in horizontal inequality due to low intergenerational mobility overall, it argues, it also is explained by the lower mobility of disadvantaged compared with advantaged groups in many societies. Group-based discrimination, among other factors, contributes. The paper further explores empirical data on this relationship across countries and shows how it may be especially problematic in the countries of the Global South.

Key words: social mobility, horizontal inequality, ethnic inequality, intergenerational mobility, discrimination

JEL classification: D63, J62, J15, J71

Acknowledgements: We are thankful for discussions with participants in the UNU-WIDER Social Mobility project, and in particular to Shahe Emran and Anirudh Krishna for detailed comments. Several key suggestions, while not addressed here in depth, are studied more fully in our further research. This paper was presented at the project workshop on ‘Social mobility in developing countries—concepts, methods, and determinants’, UNU-WIDER, Helsinki, Finland, 5–6 September 2019.
1 Introduction

Inequality matters both for normative and instrumental reasons. While most work on inequality examines inequality between individuals or households, inequality between groups in society may be as, if not more, concerning. In this paper, we focus on the latter—more specifically, on what are commonly referred to in the literature as ‘horizontal’ inequalities between ‘culturally defined’ groups (Stewart 2008: 3).\(^1\) (Inequalities between individuals or households, by contrast, are termed ‘vertical’ inequalities.)

Differences in well-being or status linked systematically to ascriptive attributes, such as skin colour, maternal language, tribal background, or indigeneity, are at odds with basic principles of equality, justice, and fairness. Moreover, research suggests significant negative implications for peace and prosperity (e.g. see Alesina et al. 2016; Baldwin and Huber 2010; Brown and Langer 2010; Stewart 2008). We know that horizontal inequalities vary across countries, as well as within countries, and change over time, yet there are major gaps in our understanding of precisely how and why (Canelas and Gisselquist 2018a, 2018b). Addressing these gaps has direct implications then not only for scholarly discussion and debate but also for policy aimed at supporting equality, social justice, and inclusion.

This paper presents a new framework for understanding the factors that influence contemporary persistence and change in horizontal inequalities. Its entry point is the relationship between horizontal inequality and intergenerational economic mobility, including consideration of why mobility may differ significantly across ethnic groups, especially in horizontally unequal societies. In so doing, it draws new links across literatures on social mobility, horizontal inequality, and ethnic politics. It draws on empirical analyses of multiple world regions, but it is anchored in consideration of the Global South.

This paper advances three interrelated arguments. First, horizontal inequality can be expected to be persistent (a) over decades and multiple generations, in situations where starting levels of horizontal inequality are high and overall intergenerational mobility low and (b) even longer, sometimes indefinitely, in situations where mobility varies across ethnic groups and is lower for disadvantaged than for advantaged ethnic groups.

Second, persistent horizontal inequality is likely a comparatively larger problem for Global South than Global North countries. Although the above-mentioned situations are present in multiple countries around the world, the extant literature suggests that average initial rates of horizontal inequality tend to be higher and overall rates of mobility comparatively lower in low and middle-income countries than in wealthier countries. Other factors, including comparative weaknesses in state capability and the rule of law, also may contribute to a greater likelihood that mobility will vary across ethnic groups and be lower for disadvantaged groups in developing countries.

Third, research suggests that it is not unusual to find variation in social mobility across ethnic groups, with comparatively lower levels for disadvantaged groups. This variation can be understood in terms of how key determinants of mobility—human capital investment and parental endowments, credit constraints, peer influences and role model effects, and other neighbourhood effects—differ across ethnic groups. The role of group-based discrimination here has been highlighted in related literature. We show how such discrimination—both current and past—is

\(^1\) For further discussion, see Section 2.
important, alongside several elements more ‘internal’ to disadvantaged groups themselves: geography and remoteness, social networks, culture, and representation. These latter elements, in particular, influence not only the ‘quality’ but also the ‘fit’ of educational and other public services—and thus the efficacy of parental investment in children.

In brief, this paper advances a new conceptual framework for understanding horizontal inequality as an outcome, hinging on the links between horizontal inequality and social mobility. The framework implies that the persistence in horizontal inequality over generations in many countries can be understood as a function of high initial degrees of horizontal inequality, low rates of social mobility overall, and variant social mobility across ethnic groups that is lowest for (certain) disadvantaged groups. Such group variation in mobility is rooted in particular in current group-based discrimination, the legacies of past discrimination, and structural and cultural characteristics of ethnic groups that influence their relationship with the state. This analysis, we argue, points to the need for state institutional reform and activist intervention to address social exclusion in a timely manner, while also implying some major constraints.

The paper is organised as follows. Section 2 gives a brief introduction to the key concepts of social mobility and horizontal inequality and their interrelations. Section 3 takes a closer look at these linkages and presents the core conceptual framework, which builds in particular on the model presented by Chetty et al. (2018). Section 4 considers the determinants of social mobility and why mobility may vary across ethnic groups, drawing on findings and examples from the extant literature. In Section 5, we turn to how horizontal inequality can be expected to change (or persist) over time given different mobility levels and assumptions. Section 6 concludes by drawing out key implications for policy and considering areas for future research.

2 Concepts and measures

Social mobility is ‘the ability to move between different levels in society or employment usually from a lower to a higher social class’.

We focus here on social mobility from one generation to the next, measured primarily in terms of income status, that is, intergenerational (income) mobility. We also consider mobility in terms of educational attainment in some of the analysis.

Much of the literature has focused on mobility in terms of intergenerational elasticity of income (IGE). A standard measure of IGE is calculated by regressing log child income on log parent income (Solon 1992; Solon et al. 1999). An alternative measure used by Chetty et al. (2014) is based on a rank–rank specification, which they have found yields (broadly comparable but) more stable estimates (see also Dahl and DeLeire 2008). In this approach, children’s incomes are ranked based on income relative to others in their birth cohort, and their parents’ incomes ranked relative to other parents with children in the same birth cohort. IGE is captured by the rank–rank slope, which identifies the association between children’s and parents’ positions in the income distribution. In the conceptual framework discussed here, we work with this latter measure.

Thus, IGE is a measure of relative mobility or positional movement (Fields, forthcoming). It speaks to the question: ‘What are the outcomes of children from low-income families relative to those of children from high-income families?’ (Chetty et al. 2014: 1560). While comparisons across countries are problematic due to differences in data and methodologies, and to the lack of longitudinal income data over generations in many developing countries, the extant literature
broadly suggests that relative mobility is highest in Nordic countries (with IGEs in the range of 0.2) and lowest in high-inequality developing countries, with other wealthy countries like the United States and Italy in between (with IGEs in the range of 0.4–0.5) (Piraino 2015: 397). For South Africa, Piraino (2015) has calculated an IGE of about 0.6. IGEs in this broad range are also estimated for instance in Chile (Nunez Javier and Miranda 2010) and urban China (Gong et al. 2012).

A useful source for cross-country comparisons, which we draw on in the following discussion, is the Global Database on Intergenerational Mobility (GDIM 2018), an effort to provide cross-country comparative information on mobility. It includes estimates of intergenerational persistence covering individuals born between 1940 and 1989 across 148 economies. Simple averages of reported data here are 0.51 for developing countries and 0.36 for high-income countries (but such averages should be interpreted with caution given challenges of comparison across countries and measures). Piraino (forthcoming) considers how differences in drivers of mobility contribute to such variation in IGEs between low-income and high-income countries.

In this paper, we also focus on absolute mobility, or the outcomes of children from families at a given income level in absolute terms. Chetty et al.’s (2018) analysis of mobility in the United States suggests the particular relevance of absolute mobility in understanding persistent inequality across racial groups. Various measures of absolute mobility are used in the literature. A simple one is the probability for a child born to a family in the bottom income quintile to earn in the top income quintile as an adult. Using data on education rather than income, another measure—based on the work by Alesina et al. (2019) on mobility in Africa—is the likelihood that children achieve higher/lower/similar educational attainment to their parents or other adults of the same generation in their household (e.g. uncles and aunts).

Considerable research attention has focused on the relationship between social mobility and vertical inequality. Becker and Tomes (1979, 1986) is a classic starting point. At the core of the model are altruistic parents who must decide how much to invest in their children. How much human capital children have in the next period is determined by parental investment in their human capital, children’s endowments, and government spending on education.

### 2.1 Horizontal inequality

Although work on social mobility has focused primarily on vertical inequality, links with horizontal inequality are also apparent. Horizontal inequality refers to inequality ‘in economic, social or political dimensions or cultural status between culturally defined groups’ (Stewart 2008: 3). We focus here on socioeconomic dimensions, for which commonly used indicators include income, consumption, and years of schooling. Various measures are used to describe horizontal inequality, ranging from simple mean differences in such indicators for different groups to more sophisticated constructs. Issues of concept and measurement are treated in some depth in the existing literature, and we do not revisit those debates here. Instead, we build on work by Mancini et al. (2008) which provides the most comprehensive consideration of these issues of which we are aware. They present a strong case for three measures: the group-weighted Gini coefficient \( G_{Gini} \), which compares the mean in the outcome variable of every group with that of every other group; the

---

3 Chetty et al. (2014) estimate IGE for the United States at 0.45. Using the other log–log measure and excluding children with zero income, they estimate IGE for the United States at 0.344.

4 In other words, we focus on two of the six commonly used concepts of mobility identified by Fields (2006).

5 For more discussion, see Canelas and Gisselquist (2018b).
group-weighted Theil index \((G_{Theil})\), which compares the mean in the outcome variable for each group with the national mean; and the group-weighted coefficient of variation \((G_{COV})\), which measures overall dispersion, capturing variation at all levels of the distribution, not only near the mean.

In this paper, we draw primarily on the often-used \(G_{Gini}\) measure (and define it more formally in Section 3) An important question for future work is whether this measure should be combined with, or even replaced by, another measure such as the \(G_{COV}\). Use of the \(G_{Gini}\) focuses our attention on variation in mean outcomes across groups rather than, for instance, in variance, which could be more informative in the study of mobility (see Emran et al. 2019).

A key challenge in analysing horizontal inequality across countries is that available data are much more limited and problematic than data on vertical inequality (Canelas and Gisselquist 2018a). In the brief analysis presented here, we draw on two data sources. The first, the Education Inequality and Conflict (EIC) dataset (see EIC 2015), offers estimates of horizontal inequality—measured using the \(G_{Gini}\)—in educational attainment (HI-E). The data used covers a sample of up to 76 countries for the year 2000 for groups classified in the EIC as ‘ethnic’, ‘religious’, and ‘wealth’ groupings. The second is Østby’s (2008) ‘horizontal economic inequality’ measure, which compares ethnic groups in 36 countries using a composite wealth index based on individual assets sourced from the demographic health surveys.

2.2 Linking social mobility and horizontal inequality

Although the relationship between social mobility and horizontal inequality is comparatively understudied and undertheorized, inequality between groups is far from absent in the literature on intergenerational income mobility. In Becker and Tomes (1979), for instance, various ethnic attributes play into the concept of endowments, which is fundamental to their model. Children’s endowments of capital are determined, it is assumed, by ‘the reputation and “connections” of their families, the contribution to the ability, race, and other characteristics of children from the genetic constitutions of their families, and the learning, skills, goals, and other “family commodities” acquired through belonging to a particular family culture’ (Becker and Tomes 1979: 1158).

One exception that we return to later is Durlauf’s memberships theory of inequality, which focuses on group-level influences on individuals (e.g. see Durlauf 2004). The literature on equality of opportunity provides another link (see Ferreira and Peragine 2015; Roemer and Trannoy 2015). Brunori et al. (2013: 17) argue that inequality of opportunity is ‘the missing link between the concepts of income inequality and social mobility; if higher inequality makes intergenerational mobility more difficult, it is likely because opportunities for economic advancement are more unequally distributed among children’.

Especially relevant to our purposes in this paper is the work by Stewart (2009), which like our analysis is centrally concerned with horizontal inequality as an outcome and the role of mobility therein (see also Stewart and Langer 2008).\(^6\) In Stewart’s (2009) argument, persistent horizontal inequality is the result of two traps. A ‘capability inequality trap’ looks at how different capabilities (education/health/nutrition) each promote other capabilities, and in turn influence productivity and income. For instance, nutritional deficiency affects a child’s educational attainment, which in turn affects their productivity and income, which in turn affects the health, nutrition, and education of their children, and so on. A ‘capital poverty trap’ addresses how access to each type of capital (namely, human, social, cultural, and physical) affects returns to other types. These traps are closely

\(^6\) Stewart (2009) specifically considers ‘group inequality’, thus including other types of groups such as gender.
related and conceived as reinforcing each other: for instance, the capabilities trap ‘leads to’ low human capital (Stewart 2009: 324). According to Stewart (2009), inequalities in capabilities and assets in turn are evident across groups, particularly due to asymmetries in social and cultural capital, as well as group-based discrimination. Stewart argues that this suggests horizontal inequalities tend to be more persistent than individual inequalities, and require special interventions to address them.

Stewart’s (2009) analysis is a useful starting point for our analysis in this paper. As we will see later, our conclusions are complementary in the sense that both analyses point towards significant forces that contribute to persistence in contemporary horizontal inequalities without special intervention.

Our analysis also differs from that of Stewart (2009) in several ways. For one, it builds much more explicitly from focused consideration of intergenerational mobility and its key determinants as identified in the literature on mobility, whereas Stewart’s is framed within a capabilities approach to human development. We make no claims that our starting point is best, but note the value for cumulating knowledge of approaching the same question from a different angle—and coming to some similar conclusions.

That said, our conceptual framework and the analysis presented here also identify—and provide stylized evidence for—some additional hypotheses about horizontal inequality as an outcome. In particular, we discuss the role of absolute and relative mobility and its variation across groups, which in turn supports expectations about the situations in which persistence in horizontal inequality is expected (and not), its drivers, and timeframes for change.

2.3 The Great Gatsby Curve revisited

Empirical studies have demonstrated a relationship between vertical inequality and intergenerational persistence in various countries—the Great Gatsby Curve, as coined by Krueger in 1992 (see Corak 2013). Typically, the measure of inequality used is the Gini-income measure, and a positive relationship with IGE has been identified: countries with high levels of income inequality have high income persistence levels across generations. Strong correlation is likewise shown between intergenerational persistence and inequality of opportunity (Brunori et al. 2013). Such empirical relationships have motivated broader theoretical consideration of underlying causal channels.

An empirical question is whether there is likewise a relationship between horizontal inequality and intergenerational persistence. Before turning to underlying channels in Section 3, we consider this question briefly using the data sources introduced earlier: GDIM (2018) on intergenerational persistence and the EIC dataset and Østby (2008) on horizontal inequality. We use three measures of intergenerational income inequality from the GDIM (2018): IGE ‘share 1’, which measures the effect of parental education on the income of their children via the effect on child education; IGE ‘share 2’, which measures the effect of parental education on the income of their children via factors other than child education; and IGE ‘share 3’, which measures the effect of parental characteristics such as income or networks on children’s income.\(^7\)

---

\(^7\) There were insufficient country data observations on relative IGM in income, which is the standard measure in the mobility literature, to plot the Great Gatsby Curve given the countries for which horizontal economic inequality data were available. The three income share measures are used as a proxy here for relative IGM in income and are derived by decomposing IGE income (for a more detailed discussion, see Narayan et al. 2018).
Figure 1 shows the relationship between mobility and the EIC’s HI-E estimates. This positive relationship obtains whether we consider HI-E for ethnic, religious, or wealth groupings. ‘Ethnic’ HI-E has the highest level of correlation with intergenerational persistence in education. Which way the causal arrow runs, however, is open to debate, and our discussion in the next sections speaks to this open question.

Figure 1: The Great Gatsby Curve for education: (a) horizontal inequality and absolute mobility; (b) horizontal inequality and relative mobility

(a) Horizontal Inequality and Absolute Mobility

(b) Horizontal Inequality and Relative Mobility

Note: the figure displays inequality measured by the GGini education coefficient (x-axis) versus absolute intergenerational mobility and intergenerational education persistence, IGP (y-axis.)

Source: authors’ compilation based on GGini from EIC (2015) and IGP from GDIM (2018).
Figure 2 shows the relationship between mobility and Østby’s (2008) horizontal economic inequality measure. Interestingly, a positive relationship can be seen between this measure and the ‘share 1’ and ‘share 2’ figures, but a negative relationship with ‘share 3’. The former is consistent with our expectation that higher levels of horizontal inequality would be associated with lower mobility, while the latter is not. Closer study of the data, including specific countries, is needed to better understand the later finding. One possibility is that higher levels of horizontal inequality are related to one group becoming better (or worse off) economically than others, leading to new circumstances that primarily affect their children’s outcome and go along with lower income persistence. On aggregate then, it may not necessarily be that the negative relationship between horizontal inequality and intergenerational persistence is progressive, but that the eventual outcome is ambiguous and should be considered on a group by group basis.

Figure 2: The Great Gatsby Curve for income

Note: the figure displays inequality measured by the GGini income coefficient (x-axis) versus intergenerational income persistence share, IGPshare (y-axis). Share 1 measures effect of parental education on the income of their children via their effect on child education; share 2 measures effect of parental education on income of the children via other factors except child education; share 3 measures effect of parental characteristics such as income or networks on the child income.

Source: authors’ compilation based on GGini from Østby (2008) and IGPshare from GDIM (2018).

3 A conceptual framework

A standard approach to considering intergenerational income persistence is with a regression-to-the-mean model (Becker and Tomes 1979):

$$\ln y_{i,t} = \alpha + \beta \ln y_{i,t-1} + \varepsilon_i$$

where $y_{i,t-1}$ and $y_{i,t}$ are the permanent incomes of the parents and children, $\varepsilon_i$ is an error term unrelated to parental income, and $\alpha$ and $\beta$ can be understood as measures of absolute mobility and relative mobility (1−IGE) as defined above.

Over the long term under simple assumptions, incomes in this model regress towards the mean—at a rate described by the IGE or $\beta$. Table 1 shows simple examples with two different values of
Consider a family that starts off in 1960 with a permanent income that is 100 per cent above the mean. If intergenerational mobility is high (e.g. $IGE=0.2$), we expect the grandchildren of this family to have incomes less than 5 per cent above the mean. If generations are 25 years and we start in 1960—for example, the year Nigeria gained independence along with 16 other Sub-Saharan African countries—this would be in 2010. If intergenerational mobility is low (e.g. $IGE=0.6$), we would expect it not to be until the seventh generation (i.e. in 2110) that the permanent incomes would be less than 5 per cent above the mean.

### Table 1: Simple mobility examples

<table>
<thead>
<tr>
<th>Generation</th>
<th>Year</th>
<th>$IGE=0.2$</th>
<th>$IGE=0.6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>1960</td>
<td>100.00%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Second</td>
<td>1985</td>
<td>20.00%</td>
<td>60.00%</td>
</tr>
<tr>
<td>Third</td>
<td>2010</td>
<td>4.00%</td>
<td>36.00%</td>
</tr>
<tr>
<td>Fourth</td>
<td>2035</td>
<td>0.80%</td>
<td>21.60%</td>
</tr>
<tr>
<td>Fifth</td>
<td>2060</td>
<td>0.16%</td>
<td>12.96%</td>
</tr>
<tr>
<td>Sixth</td>
<td>2085</td>
<td>0.03%</td>
<td>7.78%</td>
</tr>
<tr>
<td>Seventh</td>
<td>2110</td>
<td>0.01%</td>
<td>4.67%</td>
</tr>
<tr>
<td>Eighth</td>
<td>2135</td>
<td>0.00%</td>
<td>2.80%</td>
</tr>
<tr>
<td>Ninth</td>
<td>2160</td>
<td>0.00%</td>
<td>1.68%</td>
</tr>
<tr>
<td>Tenth</td>
<td>2185</td>
<td>0.00%</td>
<td>1.01%</td>
</tr>
</tbody>
</table>

Source: authors’ computation using hypothetical numbers.

Measuring income in terms of rank as defined above, Chetty et al. (2018) extend this model to incorporate racial group effects. While Chetty et al.’s (2018) model is specified for racial groups given their focus on politics of the United States, it can be straightforwardly adapted for the broader category of ‘ethnic’ groups as defined in this paper. By assuming linear conditional expectations of the mean rank of children given their parents’ rank, an individual’s income is modelled as an ethnic group-specific linear function of their parent’s income:

$$y_{i,t} = \alpha_r + \beta_r y_{i,t-1} + \epsilon_{i,t}$$

where $\alpha_r$ and $\beta_r$ are both ethnic group-specific measures of the rate of absolute and relative mobility. For simplicity, we assume they do not vary across generations.

Under the linear specification, the mean rank of individuals of each ethnic group in generation $t$ is thus:

$$\bar{y}_{r,t} = \alpha_r + \beta_r \bar{y}_{r,t-1}$$

Over the long term, Chetty et al. (2018) show that the mean rank of group $r$ converges to a steady state ($SS$):

$$\bar{y}_{r,t} = \bar{y}_{r,t-1} = \bar{y}_{r,SS} = \frac{\alpha_r}{1-\beta_r}$$

We can then consider horizontal inequality using the $GGini$ measure, which compares the mean of every group to the mean of every other group:

---

8 That is, it is not necessary ‘to track the evolution of the full distribution to characterize the evolution of mean outcomes by race’ (Chetty et al. 2018: 9).

9 In generation $t + n$, as $n \rightarrow \infty$. 

---
\[ GGini = \frac{1}{2\bar{y}} \sum_{r}^{R} \sum_{s}^{S} p_r p_s |\bar{y}_r - \bar{y}_s| \]

where \( \bar{y}_r \) is group \( r \)'s mean value, \( R \) is the number of groups, and \( p_r \) is group \( r \)'s population share.

Three simple expectations about horizontal inequality as an outcome follow from the framework above:

1. If rates of absolute and relative mobility are the same across groups, a horizontally unequal society will eventually become horizontally equal.

This follows from the above:

Over the long term, given that \( \bar{y}_{rss} = \frac{\alpha_r}{1-\beta_r} \), \( GGini_{t+n} \) becomes

\[ GGini_{t+n} = \frac{1}{2\bar{y}} \sum_{r}^{R} \sum_{s}^{S} p_r p_s \left| \frac{\alpha_r}{1-\beta_r} - \frac{\alpha_s}{1-\beta_s} \right| \]

If \( \alpha_r = \alpha_s \) and \( \beta_r = \beta_s \) for all groups, we can see that \( GGini_{t+n} = 0 \), which is perfect horizontal equality.

2. How long this takes should depend upon initial levels of horizontal inequality and rates of mobility; horizontal inequality can be persistent over multiple generations when initial horizontal inequality is high and overall mobility is low.

As above, incomes in this simple model regress towards the mean at a rate described by \( \beta \). Table 2 extends the 'low mobility' (IGE=0.6) example illustrated in Table 1 assuming a country with two ethnic groups of equal size and a constant overall mean income of 1,000. We begin in the first generation with a perfectly horizontally unequal society (\( GGini=1 \)) in which the average member of the advantaged group makes 2,000 and the average member of the disadvantaged group 0, falling to a \( GGini \) of 0.36 by the third generation, and less than 0.1 by the sixth.

<table>
<thead>
<tr>
<th>Generation</th>
<th>Year starting in 1960</th>
<th>Advantaged group (50%)</th>
<th>Disadvantaged group (50%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Income % above mean</td>
<td>Income</td>
<td>Income % above mean</td>
</tr>
<tr>
<td>First</td>
<td>1960</td>
<td>100.00</td>
<td>2,000</td>
</tr>
<tr>
<td>Second</td>
<td>1985</td>
<td>60.00</td>
<td>1,600</td>
</tr>
<tr>
<td>Third</td>
<td>2010</td>
<td>36.00</td>
<td>1,360</td>
</tr>
<tr>
<td>Fourth</td>
<td>2035</td>
<td>21.60</td>
<td>1,216</td>
</tr>
<tr>
<td>Fifth</td>
<td>2060</td>
<td>12.96</td>
<td>1,130</td>
</tr>
<tr>
<td>Sixth</td>
<td>2085</td>
<td>7.78</td>
<td>1,078</td>
</tr>
<tr>
<td>Seventh</td>
<td>2110</td>
<td>4.67</td>
<td>1,047</td>
</tr>
<tr>
<td>Eighth</td>
<td>2135</td>
<td>2.80</td>
<td>1,028</td>
</tr>
<tr>
<td>Ninth</td>
<td>2160</td>
<td>1.68</td>
<td>1,017</td>
</tr>
<tr>
<td>Tenth</td>
<td>2185</td>
<td>1.01</td>
<td>1,010</td>
</tr>
</tbody>
</table>

Source: authors’ computation using hypothetical numbers.

3. Horizontal inequality will be persistent over multiple generations—and even permanent in the steady state—when mobility (absolute and/or relative) varies across ethnic groups and is lowest for disadvantaged groups.
We illustrate this point through four scenarios in Section 5. More formal treatment can be found in Chetty et al. (2018).

4 Determinants of mobility

We predict that horizontal inequality will be persistent over multiple generations—and even permanent in the steady state—when mobility (absolute and/or relative) varies across ethnic groups and is lowest for disadvantaged groups. Section 5 draws on empirical analysis of such variation across ethnic groups in India (Asher et al. 2018; Hnatkovska et al. 2013) and the United States (Chetty et al. 2018). Other work provides evidence of such variation in mobility across ethnic groups in South Africa (Nimubona and Venkat of 2007; Piraino 2015) and Brazil (Osorio 2008). This section considers the roots of such variation and why it may be particularly apparent in the Global South.

A starting point for our analysis is Iversen et al.’s (2019) consideration of key determinants of mobility identified in the research literature. Pointing out that this work has been based heavily on study of Western countries, they explore the applicability of core findings in developing country contexts. They begin with Becker and Tomes (1979) and the influence of human capital investment and parental endowments. In developing country contexts, they note, there is evidence of significant association between parental background (especially income and education) and investment in the human capital of children, which could be seen as broadly consistent with the Becker–Tomes model. However, there are also challenges in its applicability for several reasons.

For one, the Becker–Tomes model assumes perfect capital markets, whereas the research literature points to how capital market imperfections can constrain borrowing by poor households thus restricting the ability of individuals to move into occupations requiring high initial investment. Although research on Western countries does not show strong evidence for the impact of such credit constraints on mobility, they may be more pronounced in developing countries given lack of collateral among poor households and informational constraints.

The research literature also points to the influence of neighbourhood effects on mobility (Chetty et al. 2018; Chetty et al. 2014; Durlauf 2004). Neighbourhood effects may influence mobility through several key channels. Peer influence and role model effects in particular may be especially relevant for the poor in all countries and for developing countries in general. They may affect mobility via their influence on individual aspirations; the enforcement of social norms; and the availability of information, guidance, mentorship, and connections for individuals in the pursuit of diverse educational and occupational pathways (see Krishna 2014). Such influences can be reinforcing: for instance, for the poor, there can be a ‘self-fulfilling equilibrium where low aspirations lead to low effort, which in turn reinforces low aspirations, generating persistent intergenerational inequality’ (Iversen et al. 2019: 245).

Neighbourhoods also can influence children and the efficacy of parental investment through neighbourhood institutions, especially schools (Solon 2004). Iversen et al. (2019: 245) predict neighbourhood effects to be more salient in developing countries ‘given within-country differences in the provision of public goods, in the quality of primary schooling, as well as the concentration of poverty among socially marginalized groups, who are often located in more remote areas’.

As this latter point suggests, several of Iversen et al.’s (2019) arguments with respect to low-income countries (compared with high-income countries) can be extended straightforwardly to low-income groups within these same countries. Just as mobility in low-income countries may be
hindered by poorer average quality of schooling, so too may mobility among (regionally concentrated) low-income groups within these countries be hindered by poorer average quality schooling compared with that in wealthier communities. However, the relationship between mobility and ethnic groups can also be more complex. In the next section, we turn to ethnic groups as a concept and unit of analysis, which lays the basis for our consideration of the factors identified in Iversen et al. (2019) in the context of ethnic groups.

4.1 A constructivist approach to ethnic groups as a concept and unit of analysis

While popular discussion regularly treats an individual’s ethnicity as a simple ‘fact’, it is not so straightforward for scholars who study ethnicity; defining and conceptualizing the ‘ethnic group’ has been the subject of significant debate in the research literature (Bates 2006; Brubaker 2004; Chandra and Wilkinson 2008; Hale 2004; Varshney 2007). On one side are approaches suggesting deeper roots and greater fixity. Smith (1986), for instance, identifies key characteristics of ethnic groups to include a common myth of ancestry, sense of common history, shared culture, attachment to a specific territory, and a sense of solidarity.

On the other are approaches highlighting ethnic identities that are more fluid, socially constructed, situationally contingent, and (sometimes) instrumental. In this latter approach, ‘ethnic group’ is a broad concept comprising those linked with a diverse set of ascriptive attributes inherited at birth, including skin colour, maternal language, tribe, caste, religion, and sometimes region (Chandra 2004; Horowitz 1985). Scholars adopting this latter approach describe all of the following as ‘ethnic groups’: Hindus and Muslims (Varshney 2007) and scheduled castes in India (Chandra 2004); Africans, whites, ‘coloureds’, and Indians in South Africa (see Ferree 2010); Bemba, Nyanja, Tonga, and Lozi speakers in Zambia (Posner 1998); and indigenous and non-indigenous populations in Latin America (Htun 2004; Van Cott 2007). Thus, the sharp conceptual distinctions highlighted in some other work among, for instance, ‘ethnic’, ‘linguistic’, and ‘religious’ divisions (Alesina et al. 2003) or ‘ethnic’ and ‘racial’ groups are not central in this approach.11

The approach taken here falls broadly within this latter constructivist/instrumentalist camp. This is now the dominant one in comparative research on ethnic politics, which has clearly shown that ethnic groups are not immutable and has rejected the simple ‘primordialist’ position (Chandra 2001). However, we deviate from the widely used definition above, which can be understood to strip all content from ethnicity, making it difficult to understand the conceptual and effective power of ethnic groups as distinct from any other types of group or category (gender, social class, occupation, location of residence). We build here instead on Fearon (2003) in defining ethnic groups in terms of prototypical features, some of which echo Smith (1986). These include membership that is ‘reckoned primarily by descent by both members and nonmembers’; members deriving normative and psychological value from membership; sharing of some cultural features, such as language, religion, and customs, by the majority of group members; having a homeland or memory of one; and a sense of shared collective history (Fearon 2003: 201). While such features are prototypical, an actual ethnic group might lack some of them (see also Fearon and Laitin 2000).

In summary, ethnic groups are social constructions defined by both members and non-members and these social constructions in turn can influence social interactions. Both the external and internal elements here are important in understanding why and how mobility may vary across

10 We use ‘group’ loosely here to refer not only to collectivities that are organized and mobilized but also to those that may be more like categories (see Brubaker 2004).

11 Indeed, it raises questions about the conceptual rigour with which such distinctions can be made.
ethnic groups. We begin with external elements: how non-members treat members of ethnic
groups, focusing on group-based discrimination. We then turn to some key internal elements—
what commonalities are shared by group members—and how they may influence mobility at the
group level.

4.2 Group-based discrimination

Research on social mobility and ethnic groups has paid particular attention to group-based
discrimination and bias. Within the broad framework of the Becker–Tomes model, it can be
understood to influence the efficacy of human capital investment by disadvantaged-group parents
via multiple channels outlined above. We draw here primarily on studies of the United States where
the evidence base on discrimination across multiple sectors is most developed, offering selected
examples from studies in developing countries.

Discrimination in the labour market implies that the same level of investment by parents
corresponds to lower outcomes for children from disadvantaged, discriminated-against groups. In
the United States, research over several decades has found evidence of racial disparities and
discrimination in the labour market (Bendick Jr. 1998), including its role in the underrepresentation
of African Americans in managerial, sales, and other occupations (Gill 1989). While some racial
disparities in interviewing, hiring, and advancement can conceivably be attributed to other job-
relevant candidate characteristics, experimental studies offer evidence for the isolated effects of
ethnic bias and discrimination. In a field experiment in Boston and Chicago involving random
assignment of white and African American names on resumes, Bertrand and Mullainathan (2004)
show significant racial differences in callback rates for job applications and market rewards for
having a better resume. To get one callback, applicants with white names sent about 10 resumes,
while those with African American names sent about 15 resumes. This gap between white and
African American applicants widens with resume quality; higher quality resumes have a smaller
effect on callbacks for African American compared with white applicants. Similarly, in a field
experiment in New York City, Pager et al. (2009) find significant differences in callbacks and job
offers for entry-level jobs for white, black, and Latino job applicants with equivalent resumes.
Black applicants fared about half as well as equally qualified white applicants. White applicants just
released from prison fared about as well as black and Latino applicants without prison records.
Using similar resume audit experiments, Galarza and Yamada (2014) demonstrate ethnic
discrimination against indigenous compared with white job applicants in Lima, Peru, while,
focusing on software jobs in Delhi, India, Banerjee et al. (2009: 14) find somewhat conflicting
evidence on caste discrimination, concluding that, overall, ‘caste identities do not significantly
affect the callback decisions of firms’. However, other significant work on India documents
widespread patterns of caste discrimination (Deshpande 2011; Thorat and Newman 2012).

Discrimination in education likewise can imply that the same level of investment by disadvantaged-
group parents corresponds to lower outcomes for their children. Research in the United States
finds disparities both in how teachers discipline children of different ethnicities for equivalent
behaviour (Downey and Priaples 2004; Okonofua et al. 2016), as well as in their expectations for
children of different ethnicities (Harber et al. 2012; Tenenbaum and Ruck 2007), which may
negatively influence educational outcomes for Latino and African American children in particular
compared with white children. Evidence suggests that discrimination and implicit bias even

12 Selected studies show variation across ethnic minority groups. Drawing on four quantitative meta-analyses,
Tenenbaum and Ruck (2007) report evidence that teachers actually hold highest expectations for Asian American
children, followed by children of European descent.
influence children at the preschool level, especially for black boys, consistent with disproportionate rates of preschool expulsion (Gilliam et al. 2016).

Discrimination may also influence mobility via credit constraints for members of disadvantaged groups. As Weller (2009: 7) summarizes: ‘Household debt can be, if used correctly, the grease for economic mobility. By borrowing, many more families can afford to buy a home, car, or a college education than would otherwise be the case. And debt allows families to smooth out income fluctuations due to short-term spells of unemployment, a medical emergency, among others.’ Indeed, access to credit is ‘vital’ for members of disadvantaged groups ‘if they are to overcome the low level of their initial endowments’ (Munnell et al. 1996: 25).

Research in the United States shows minorities are more likely than whites to be credit constrained (Duca and Rosenthal 1993). With respect to mortgage loans, for instance, Avery et al. (1993) find ethnic (racial) disparities in approval rates controlling for applicant income, region, and other factors, while Munnell et al. (1996) show minority applicants in Boston denied a mortgage almost twice as often as white applicants with similar observable default risk factors. Minorities may also receive less favourable loan terms and may be encouraged to apply for less desirable loans (see Ross and Yinger 2002). Likewise, controlling for differences in creditworthiness and other factors, Blanchflower et al. (2003) show black-owned small businesses about twice as likely to be denied credit.

Discrimination by mortgage-lending institutions is one component of discrimination in housing markets. In the United States, some of the strongest evidence for discrimination is provided by a series of studies conducted since the late 1970s by the Urban Institute, sponsored by the United States Department of Housing and Urban Development. These studies have employed a paired-testing or audit methodology in which two people, one white and one minority, pose as equally qualified home-seekers (Turner and James 2015). With some changes over years, these studies document that African American and Hispanic home-seekers receive less favourable treatment than white home-seekers, receiving less information and being shown fewer homes and apartments (Turner et al. 2002; Turner et al. 2013).

Finally, there is evidence of discrimination influencing the distribution of public resources and services, although it is difficult to isolate whether it is ethnic discrimination at work or lower public resources in poorer neighbourhoods or discrimination based on economic status or class. In Chicago, Mladenka (1989: 51) finds that ‘class has replaced race as the primary determinant of the service distribution pattern’. Using an email correspondence study posing simple queries to local public service providers, Giulietti et al. (2017) find better response rates for emailers with white-sounding names than for those with black-sounding names. In a broadly similar national field experiment in China, Distelhorst and Hou (2014) find local officials 33 per cent less likely to provide assistance to electronic communication from citizens with Muslim names than to those with ethnically-unmarked names.

4.3 Parsing group-based discrimination

Discrimination is used as an umbrella term for several different phenomena. Parsing what ‘discrimination’ means thus can help us to more precisely specify the underlying mechanisms through which discrimination affects mobility, and what we expect to observe in developing versus developed countries. At least three distinctions are key for our purposes.

The first distinction links discrimination with formal or informal institutions or practice (e.g. de jure or de facto law). In general, we expect informal institutions and practice to be the larger and more persistent challenge today. The studies of contemporary United States experience referenced
earlier are suggestive of this point: despite various anti-discrimination legislation, such as the Fair Housing Act, discrimination persists through the significant role of informal institutions. At the international level, although not always respected, international norms and law—such as International Covenant on Civil and Political Rights or the International Convention on All Forms of Racial Discrimination—also place some clear limits with respect to explicitly discriminatory legislation.

That said, discrimination in formal institutions is also clearly at work in some contexts. Citizenship laws, for instance, can effectively exclude particular groups from rights to national public resources and services, and are an area in which more research is needed (see Manby 2018). Discrimination against particular groups need not be done by explicitly naming them, but also by defining the members of the nation in such a way that some groups are excluded. For instance, Myanmar’s 2008 Constitution and 1982 Citizenship Law define citizenship in terms of officially recognized ‘national races’, effectively excluding from citizenship Rohingya and members of other ethnic groups not included in the list (Cheesman 2017).

A second distinction to highlight, particularly in considering discrimination in informal institutions and practice, is between intentional and unintentional discrimination. For instance, some public officials may purposively provide services or carry out their duties with the clear intention of discriminating against ethnic groups they do not like. Others may have biases against disadvantaged groups that affect the way they interact and provide services, even when they are not consciously aiming to be discriminatory. Such implicit biases, for instance, are shown in the study of preschool educators cited earlier (Gilliam et al. 2016).

Unintentional discrimination can also stem from institutional weaknesses or financial constraints that hinder the implementation of the law. For instance, all children may have a legal right to primary schooling within a reasonable commuting distance from their homes, but this may be challenging and expensive to implement in remote, sparsely populated rural areas and where ethnic minorities are concentrated.

Third is the distinction between current and historical discrimination. Many of the studies above highlight ongoing discrimination. By contrast, Durlauf (2006: 151) in particular argues that ethnic differences in socioeconomic attainment are due much more to the ‘lasting effects’ of a history of discrimination via neighbourhood effects than to ongoing discrimination. Our review of the literature is more consistent with the influence of both current and historical discrimination, alongside other ethnic factors.

The preceding discussion draws heavily on the United States case, for which the evidence base across multiple areas is strongest. We expect discrimination to be at least as problematic for mobility in the Global South for several reasons. For one, the rule of law in developing countries tends to be weaker than in developed countries. According to Pritchett (2017: 24), the ‘deals space’—or ‘range of informal and personalized relationships that are observed between economic actors and political elites’—tends to be bigger in developing countries. This allows more space for discrimination in personalized informal institutions and practices. Weaker judicial systems overall also suggest more limited possibilities of legal recourse when anti-discriminatory laws are violated (see Hyden et al. 2004). Moreover, resource constraints in low-income countries may exacerbate challenges of weak capability.

In addition, as we consider later in the discussion, there is some evidence to suggest that ethnic divisions are at least—and possibly more—apparent in the Global South on average than in the Global North. Cross-country comparisons, for instance, show standard rates of ‘ethnic fractionalization’ to be highest in developing countries and Sub-Saharan Africa in particular
compared with other regions (see Alesina et al. 2003). Research arguably suggests ethnic political parties, ethnic conflict, and ethnic balancing are comparatively more important in understanding contemporary politics of ‘Southern’ regions—especially Sub-Saharan Africa—than ‘Northern’ regions, although there remains considerable debate on this point (e.g. Basedau and Stroh 2012). If ethnic identities are indeed more socially, politically, and economically ‘salient’ in developing countries than in developed countries as these various literatures suggest, we would expect ethnic group-based discrimination and biases to play at least as large a role in developing countries as in developed countries.

4.4 Other ethnic factors

Even without discrimination, typical characteristics of ethnic groups also may contribute to differential rates of mobility across groups in horizontally unequal societies. We consider four such characteristics here. Alongside group-based discrimination, geographic factors have been a focus in the extant literature on mobility. Building more explicitly on the literature on ethnic politics, we then briefly consider three additionally important factors: social networks, culture, and representation.

Residential patterns, geography, and remoteness

Many (although not all) ethnic groups are linked to particular regions. They may have a historic ‘homeland’ and be regionally concentrated in and near that area. Regional concentration also can be a result of state policies—for instance, the forced relocation of black South Africans to Bantustans as part of the policy of apartheid. Ethnically segregated residential patterns may likewise stem from informal ethnic discrimination in housing markets, as well as through migration patterns as families relocate to be closer to each other, or industries recruit via networks disproportionately from particular ethnic communities.

Neighbourhood effects are the clearest channel through which ethnic geography affects differential rates of mobility. For one, they are seen to influence the quality (and quantity) of public services. This may be due to discrimination as well as to the fact that poor communities have fewer resources to support quality public services (Durlauf 2006). State capacity may also be at play when disadvantaged minority groups live in remote areas, where the state’s reach is weakest. Likewise, geography can link with peer influence and role models in the ways outlined above.

Social networks

Ethnicity can be an important factor in social organization. Residential segregation is sometimes a contributing factor. For instance, indigenous groups residing in a remote area may simply have little opportunity to interact with non-co-ethnics. However, residential segregation and opportunity are not the only factors explaining the density of ethnic ties; we also find social networks marking ethnic membership for individuals not living in ethnically segregated communities. For instance, McDoом’s (2019) analysis of intermarriage between Muslims and non-Muslims in Mindanao, the Philippines, suggests that opportunity given residential segregation is only one of multiple factors influencing cross-ethnic relations. Not only may ethnic ties be denser, they may also be deeper or stronger than non-ethnic ties, and better at enforcing social norms (see Fearon and Laitin 1996).

Ethnic social networks thus can influence differential rates of mobility through neighbourhood effects, as well as peer influence and role models. Even without ethnically segregated neighbourhoods, we would expect influence on the latter.
An empirical question is whether ethnicity is more important in social networks in the Global South compared with the Global North. Indeed, the movement in social organization from \textit{Gemeinschaft} to \textit{Gesellschaft}—from group identity to the rational marketplace (Tönnies 1957 [1887]), from Weber’s ‘community’ to ‘society’—as countries develop has been a core component in theories of modernization. Classic modernization theory suggested a sort of one-way shift (e.g. Lipset 1960), while modernization theory critics pointed out that ethnic attachments remain salient in many industrialized societies. An important line of work demonstrates linkages between processes of modernization and the salience of ethnic networks (Bates 1974; Hechter 1974).

\textit{Culture}

Members of ethnic groups typically share some common cultural features, such as a common language, religion, customs, or social norms.\footnote{This is not to say that cultural boundaries are static; indeed, in Hechter’s (1974) argument, cultural difference itself emerges from unequal relations between the core and periphery, while cultural difference and geographic segregation likewise may be mutually reinforcing (see Barth 1969).} Because such cultural factors may contribute to making communication easier within rather than across ethnic lines (Deutsch 1966; Hardin 1995), one channel through which they influence mobility is through their relationship with social networks along the lines discussed above.

A second channel is the influence of culture on the efficacy of human capital investment given diverse preferences across groups regarding public goods. For instance, members of different groups may have distinct preferences regarding the primary language of instruction in schools, or the primary language of government. This in turn may have implications for the ‘fit’ of public services provided for majority versus minority groups. Alternatively, in analysis of the sources of ethnic inequality in Viet Nam, van de Walle and Gunewardena (2001: 1) point both to geography—that is, the fact that minorities tend to be concentrated in ‘upland and mountainous areas, often with worse access to public services and lacking basic infrastructure’—and to a sort of culture of poverty and different behavioural patterns between majority and minority groups. The latter, they suggest, may emerge as compensating behaviours and behavioural responses to their disadvantaged status.

In a still broader sense, public institutions generally can be seen to reflect the cultural dominance and preferences of the dominant or majority group, with implications in wide-ranging areas (Gisselquist 2019a). For instance, ‘justice’ as embodied in many legal systems in Latin America is rooted in Western law traditions that differ in some ways substantially from those of indigenous communities. The challenge then, Brinks (2019: 348–9) argues, is not only to improve ‘access’ to the existing justice system, but to address ‘inequality within the system itself’ by ‘equalising the conditions under which they can shape the landscape and contest the outcomes’, and to pursue alternative notions of substantive justice ‘that will more closely reflect their own normative framework’.

Likewise, notions of ‘cultural capital’ can be closely linked to the dominant group culture to the detriment of minority cultures. Farkas et al. (1990), for instance, find that ‘cultural capital’ can influence teachers’ judgement of students’ non-cognitive and cognitive performance, and also student performance.
We have seen the importance of peer influence and role model effects to mobility. The salience of ethnicity in many societies can add a further dimension: it may be that in-group peer and role model effects—from individuals of the same ethnic group—matter most for children (see Yancey et al. 2002). Co-ethnicity may also influence figures in authority in their evaluations of children’s performance and behaviour (see Dee 2005).

In horizontally unequal societies, adults from disadvantaged ethnic groups are less likely to have high socioeconomic status compared with those from dominant ethnic groups, implying simply fewer high-status in-group role models for children from disadvantaged groups. In short, for disadvantaged groups, representation—in high-status occupations and positions of influence—may matter for mobility. For members of small minorities, the numbers of such role models ‘like them’ can be small, even proportional to the group’s population share.

5 Implications of variation in mobility for horizontal inequality: some scenarios

The preceding discussion presents a framework for understanding the persistence of horizontal inequality. Building on this framework, we now take a closer look at what persistence means. We begin with three scenarios that take into account the impact of absolute and relative mobility on horizontal inequality under simple assumptions. In these three scenarios, we assume that mobility is constant over generations. We then relax this assumption in the final scenario to consider how mobility might change over generations.

5.1 Scenario 1—absolute and relative mobility are constant across groups

Suppose we could make absolute and relative mobility constant across ethnic groups. In other words, the predicted mean income rank of children, given that their parents are of the lowest rank, does not vary across groups, nor does IGE. Consider three IGE levels: 0.2 (e.g. Sweden), 0.4 (e.g. the United States), and 0.6 (e.g. South Africa).

In the first generation, given two groups, horizontal inequality would imply that one group has higher income levels than the other and a fraction of these disparities are inherited by the next generation. Figure 3 shows that by the third generation, individuals in countries with lower levels of persistence (high mobility) do not inherit any benefits. Given medium level persistence, as in the United States, it takes four generations for disparities to be eliminated, and in high persistence countries, up to five generations.
5.2 Scenario 2—absolute mobility is constant across groups, relative mobility varies

We now consider a scenario in which absolute mobility is constant across groups, but relative mobility differs. This scenario assumes that the child’s rank, given that the parent’s household is of the lowest rank, is constant across all groups, but intergenerational persistence varies across groups. Using relative mobility data for the groups from Chetty et al. (2018), we utilize hypothetical absolute mobility data to explore the implications for horizontal inequality.

At lower levels of absolute mobility, where the child’s rank is low given that their parent’s household is also of the lowest rank, Figure 4a shows that the steady-state intergenerational gap (where the parent’s household income rank is equal to the child’s rank) between the groups is modest at less than 5 percentile difference between the best-off and worst-off groups. As absolute mobility levels increase, the steady-state gap between groups rises. At high levels of absolute mobility where children can attain a greater rank given that their parents are of the lowest rank, as illustrated in Figure 4c, we observe a sizeable steady-state intergenerational gap between the best-off and worst-off groups of approximately 20 percentage points. This implies that initial levels of absolute mobility have implications for horizontal inequality, with higher initial levels associated with greater horizontal inequality in the steady state.
Figure 4: Group mobility projections with constant absolute mobility rates, varying relative mobility

5.3 Scenario 3—absolute mobility varies across groups, relative mobility is constant

Finally, we assume that relative mobility is constant across groups, but absolute mobility varies, as we would expect to see in a horizontally unequal society. We propose a stylized example using absolute mobility data for the five groups from Chetty et al. (2018) and relative elasticity in income (IGEs) for the three countries from Scenario 1.

Here, we see that in the steady state (where the group curve intersects the 45-degree line), the intergenerational gap between groups is lower in countries that have lower intergenerational persistence (see Figure 5). Using Swedish relative mobility rates (IGE=0.2), we see that the steady state for Groups 2 and 5 is at the 30th percentile while that of Group 2 (which has the highest mean child rank) is roughly at the 60th percentile. This implies an intergenerational gap of about 30 per cent. For the medium persistence level in our stylized example (IGE=0.4), the steady state for Groups 2 and 5 is attained at the 42nd percentile while Group 2’s is at the 80th percentile, implying an intergenerational gap of about 40 per cent. If persistence is very high, the steady state is attained at higher levels of parental income rank and the intergenerational gap between the most deprived group and the best-off group becomes much larger.
5.4 Scenario 4—relaxing the assumption that mobility is constant over time

In what follows, we extend Scenario 3 and relax the assumption that mobility is constant over time. We adopt the rank–rank specification of intergenerational persistence in income as set out in Section 2 and focus on absolute upward mobility and relative mobility. We use what is referred to by Stuhler (2014) as ‘extrapolation by exponentiation’, along with empirical data for different racial groups in the United States and caste groups in India. As Table 3 suggests, in this scenario, differences in mobility between the groups should be eliminated by the third to fourth generation, in line with Becker and Tomes’ (1986) expectations.

Table 3: Absolute and relative mobility

<table>
<thead>
<tr>
<th>Country</th>
<th>Generation</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
<th>Group 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States*</td>
<td>Absolute mobility</td>
<td>Percentile attained</td>
<td>36.82</td>
<td>25.43</td>
<td>36.14</td>
<td>51.44</td>
</tr>
<tr>
<td></td>
<td>Relative mobility</td>
<td>Generation 1</td>
<td>0.32</td>
<td>0.28</td>
<td>0.18</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Generation 2</td>
<td>0.102</td>
<td>0.078</td>
<td>0.032</td>
<td>0.068</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Generation 3</td>
<td>0.033</td>
<td>0.022</td>
<td>0.006</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Generation 4</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>India**</td>
<td>Absolute mobility</td>
<td>Percentile attained</td>
<td>46.82</td>
<td>46.82</td>
<td>44.23</td>
<td>4.423</td>
</tr>
<tr>
<td></td>
<td>Relative mobility</td>
<td>Generation 1</td>
<td>0.219</td>
<td>0.219</td>
<td>0.196</td>
<td>0.196</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Generation 2</td>
<td>0.103</td>
<td>0.103</td>
<td>0.087</td>
<td>0.087</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Generation 3</td>
<td>0.011</td>
<td>0.011</td>
<td>0.007</td>
<td>0.007</td>
</tr>
</tbody>
</table>

Source: authors’ computation based on empirical data from Chetty et al. (2018) for the United States; Asher et al. (2018) for absolute mobility in India; and Hnatkovska et al. (2013) for relative mobility in India (which takes scheduled tribes and scheduled castes to be in the same group as they are both protected entities in the Indian Constitution). Assume that zero persistence in income attainment is achieved when relative mobility equals approximately less than 1 per cent (<0.001).
In these data, the levels of relative mobility, which vary across groups, are shown to be much higher in India than in the United States. After four generations in both countries, advantages from the grandfather to the great-grandchildren are wiped out and what remains is the intergenerational gap or the absolute inequality measure. Using empirical data from India and assuming that absolute mobility remains constant, what this means is after four generations, relative mobility rates are equal between the groups, but the children of Muslims only attain the 29th percentile given that their parents are of low rank. This is in comparison to the children of the forward tribes who are in the 41st percentile.

Of course, our analysis recognizes the weakness of using the extrapolation methodology. The geometric decline in persistence has been argued to be unfounded by various researchers (Lindahl et al. 2015; Solon 2014; Stuhler 2014). However, empirical studies do not have a consensus on actual persistence levels across generations. Solon (2014) extends the Becker–Tomes model to include grandparents and theoretically shows a negative coefficient from grandparents. This implies that the autocorrelations decline more rapidly than is given using the geometric analysis. He argues that this is unfounded given the empirical data (such as in Clark 2012; Clark and Cummins 2014; Lindahl et al. 2015; Stuhler 2014), and provides further extensions that account for positive higher-order autocorrelations and for slower decline in persistence than can be seen geometrically. However, the data requirements to adequately capture intergenerational mobility, particularly for income, across multiple generations are more severe and, given the absence of definitive consensus on the matter, the geometric methodology remains the best alternative way to predict intergenerational persistence.

6 Conclusion and extensions

Although a growing literature examines horizontal inequality, much more attention has been paid to its implications than to its determinants. To the extent that scholarly research has considered horizontal inequality as an outcome, it has focused on long-ago origins due in particular to colonial histories and geographic factors (Alesina et al. 2016; Horowitz 2000; Michalopoulos and Papaioannou 2013) with insufficient attention to contemporary persistence and change (Canelas and Gisselquist 2018b). Addressing this gap in the research literature has direct implications for both theory and policy. This paper presents a new framework for understanding the factors that influence contemporary persistence and change in horizontal inequalities based on consideration of the relationship between intergenerational economic mobility and horizontal inequality. The discussion draws new connections between work on horizontal inequality, intergenerational economic mobility, and ethnic politics, while the simple model presented builds in particular on Chetty et al. (2018).

Several key points follow from this framework. First, long-term persistence in horizontal inequality can be understood as a function of high initial rates of horizontal inequality, low rates of social mobility overall, and/or variant social mobility across ethnic groups that is lowest for (certain) disadvantaged groups. High initial horizontal inequality or/and variant rates of relative mobility furthermore can mean horizontal inequality that persists indefinitely, even in a society exhibiting overall mobility and growth. Second, such persistent horizontal inequality is likely a comparatively larger problem for Southern countries than for Northern countries. And third, the roots of such variant mobility across groups can be seen both in how other groups treat and have treated disadvantaged groups (i.e. group-based discrimination) and in the features of ethnic structure and the groups themselves (ethnic geography, social networks, culture, and representation).
6.1 Levers of change

For those interested in influencing positive change in horizontal inequality (i.e. supporting social inclusion), this paper points to several key issues.

First, horizontal inequality can be expected to be very persistent in many contexts without policy intervention (or another big exogenous shock). Policy interventions can be aimed towards a handful of broad aims: (1) improving intergenerational mobility overall; (2) equalizing relative mobility across ethnic groups, especially bringing rates for disadvantaged groups up to that of other groups; (3) improving relative mobility for disadvantaged groups above that of advantaged groups; and (4) addressing absolute inequality and horizontal inequalities in parental endowments.

The first and second aims are much in line with universalist approaches to social policy, that is, improving policies overall and ensuring that coverage is truly universal. They are consistent with the language of ‘equal opportunity’. The third and fourth aims suggest more targeted and activist intervention. They are consistent with the language of ‘positive discrimination’ for disadvantaged groups.

Second, universalist approaches can be effective in some situations, but deep horizontal inequality will require targeted approaches and positive discrimination. If initial levels of horizontal inequality and ethnic differences in absolute mobility are significant, horizontal equality cannot be achieved through universalist approaches alone. Even in best-case scenarios, universalist approaches can take generations for horizontal equality to be achieved.

When relative mobility is constant or low, horizontal inequality persists across generations because of the intergenerational gap, which is the difference in absolute mobility between groups. What could be effective in this scenario are targeted policies that reduce disparities between groups within the generation. This could be in the form of cash transfers, wage subsidies for employing disadvantaged groups, tax credits for individuals at the lower end of the wage scale, or efforts to reduce segregation between groups in neighbourhoods and schools (on policy drivers to increase intergenerational mobility, see Narayan et al. 2018). However, when relative mobility is high, there is need to focus on reducing inequality across generations. For example, targeted policies such as cash transfers may reduce the gap in income within a generation but unless the intervention also targets the children’s outcomes conditional on the parent’s income rank, then the intervention is unlikely to have an effect in the long term. Therefore, to effectively combat this would require interventions that target both mobility and horizontal inequality at the same time (see Chetty et al. 2018).

Third, time scale matters and has implications for politics. On the one hand, given that horizontally inequality can go along with a risk of conflict and instability (United Nations and World Bank 2018), persistent horizontal inequality and slow-moving approaches can pose political risk. On the other hand, targeted interventions that may be necessary to speed up—or make possible—positive change are also risky and can promote instability. They can be deeply controversial and linked in some cases to backlash by other groups and increased ethnic tensions (see Brown et al. 2012).

Finally, the above discussion has implications for some particular levers of change depending on the factors influencing ethnic mobility levels. For instance, access to justice and legal empowerment initiatives may have some success in addressing discrimination linked to informal institutions and practice by informing individuals from disadvantaged groups of their legal rights and providing assistance in the exercise of those rights (Gisselquist 2019a; Goodwin and Maru 2017). Overall, this paper suggests that the role of the state is important in addressing horizontal inequality through mobility. In particular, attention needs to be paid to both the ‘quality’ and the
‘fit’ of public services for disadvantaged groups. For instance, attention should be paid to ensuring that schools serving disadvantaged populations have sufficient teaching materials and well-trained teachers comparable to those serving advantaged group populations; consideration also should be given to issues of ‘fit’ such as the language of instruction and the cultural biases that may be implicit in textbooks and other teaching materials.

6.2 Implications for future research

This paper points to at least three core areas for future research. The first relates to extensions of the framework presented in this paper and testing of its predictions against empirical data. The broad framework could be developed and extended in several directions—for instance, as suggested in Section 5, for more realistic assumptions about the evolution of mobility over time. Likewise, as this paper presents stylized scenarios, further work could draw more on real data, testing predictions derived from the framework against these data on selected developing countries.

A second core area for future research concerns why some horizontal inequalities have remained locked-in over generations, while others have been more fluid. The literature has explored various factors contributing to persistence in horizontal inequalities (see Stewart and Langer 2008), as well as various factors that help to explain their origins, including colonialism and conquest (Horowitz 1985; Stewart and Langer 2008), historical institutions (Michalopoulos and Papaioannou 2013), and geography (Alesina et al. 2016; Michalopoulos 2012). But very little research on horizontal inequality has comparatively considered how horizontal inequalities between different types of groups (e.g. migrants and ‘natives’ compared with racial groups) may evolve differently over time (see Gisselquist 2019b). In this context, more attention is needed in this body of work to understand how variation in rates of mobility across different disadvantaged minority groups might contribute to different evolution in horizontal inequalities.

A final core area for future work relates to levers of change and paths to reform—including tensions and policy trade-offs—as suggested in the discussion above. There are number of useful studies upon which to build, many (but not all) focused on the experiences of particular countries (e.g. Adam 1997; Deshpande 2013; Weisskopf 2004). Further building of knowledge in this area is important in informing policymaking to support more equal and inclusive societies.

References


