NBER WORKING PAPER SERIES

UNDERSTANDING THE HETEROGENEITY OF INTERGENERATIONAL MOBILITY ACROSS NEIGHBORHOODS

Neil A. Cholli
Steven N. Durlauf
Rasmus Landersø
Salvador Navarro

Working Paper 33035
http://www.nber.org/papers/w33035

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2024

Cholli thanks the National Science Foundation (NSF) Graduate Research Fellowship Program (Grant Number DGE-1746045), the U.S. Department of Education (ED) Institute of Education Sciences (IES) (Grant Number R305B140048), the Stone Center for Researchon Wealth Inequality and Mobility at the University of Chicago, and the Klarman Fellowship at Cornell University for financial support. Durlauf thanks the James M. and Cathleen D. Stone Foundation for generous support. Navarro thanks the Social Sciences and Humanities Research Council of Canada for financial support. The views expressed herein are those of the authors and do not reflect the views of the NSF, IES, ED, or National Bureau of Economic Research.

Understanding the Heterogeneity of Intergenerational Mobility across Neighborhoods
Neil A. Cholli, Steven N. Durlauf, Rasmus Landersø, and Salvador Navarro
NBER Working Paper No. 33035
October 2024
JEL No. D30, H0, J0, R0

ABSTRACT

Recent research has uncovered large spatial heterogeneity in intergenerational mobility across neighborhoods in countries around the world. Yet there is little consensus on the reasons why mobility is high in some neighborhoods and low in others. This paper analyzes a generalized mobility model that examines the roles that families' selection into neighborhoods and locational characteristics play in generating this spatial heterogeneity. We use administrative data from Denmark to decompose variation in mobility across nearly 300 larger and 2,000 smaller neighborhoods along these dimensions, accounting for sampling error. Families' selection into neighborhoods and sampling error explain most observed heterogeneity across neighborhoods. Our generalized model explains most of the differences in mobility between neighborhoods, though a small but persistent difference remains between neighborhoods that our model cannot account for. An analysis of this "irreducible heterogeneity" suggests that neighborhoods exhibit multiple types in terms of their mobility effects.

Neil A. Cholli
Cornell University
109 Tower Rd.
Ithaca, NY 14853
nac85@cornell.edu

Steven N. Durlauf
University of Chicago
Harris School of Public Policy
1307 E. 60th Street
Chicago, IL 60637
and NBER
durlauf@gmail.com

Rasmus Landersø
Rockwool Foundation Research Unit
Ny Kongensgade 6
Copenhagen 1472
Danmark
rl@rff.dk

Salvador Navarro
Department of Economics
University of Western Ontario
Social Science Centre, Room 4071
London, Ontario, Canada, N6A 5C2
snavarr@uwo.ca

A data appendix is available at http://www.nber.org/data-appendix/w33035

1 Introduction

Over the last decade, a growing body of empirical evidence has documented that intergenerational income mobility varies across neighborhoods—in other words, that children from the same parental income background have vastly different predicted later-life income based on their neighborhood of residence during childhood. Although there is a general agreement among scholars that different neighborhoods possess varying levels of mobility, there is no consensus on the reasons behind why some neighborhoods have high mobility while others have low mobility. Understanding the sources of heterogeneity in neighborhood mobility is the focus of this paper.

We employ rich administrative data to estimate a generalized model of intergenerational income mobility. First, we decompose our neighborhood mobility estimates into two underlying components: selection and location effects. By "selection," we refer to the idea that the composition of observed and unobserved (to the analyst) family characteristics within a neighborhood can produce heterogeneity in neighborhood-level mobility when an analyst does not condition on these factors. In contrast, "location effects" refer to the remaining neighborhood mobility that is not explained by selection. We examine the extent to which these effects vary across neighborhoods and their relative contributions to neighborhood mobility, accounting for estimation error. Next, we analyze which location characteristics (e.g., urbanicity) explain heterogeneity in location effects purged of selection. By doing so, we are able to isolate what we call "irreducible heterogeneity"—variation in the remaining residual mobility parameters that cannot be explained by either selection or location characteristics. We then examine this irreducible heterogeneity to determine if it reveals any discernible patterns. Here we find evidence of distinct types of neighborhoods.

Formally, our work represents a generalization of the standard, workhorse model of neighborhood mobility,

$$Y_i^c = \alpha_n + \beta_n Y_i^p + \varepsilon_i, \tag{1}$$

where i indexes family dynasties, n indexes neighborhoods, and Y_i^p and Y_i^c are, respectively, measures of parental income and child income measured in adulthood (e.g., log levels, ranks within national income distributions). The parameters α_n and β_n denote neighborhood-specific intercepts and slopes, respectively, which form the basis of popular measures of neighborhood mobility such as the intergenerational elasticity of income (IGE) or absolute upward mobility (AUM). To gather preliminary evidence regarding the sources of heterogeneity in mobility across neighborhoods, a common approach in empirical research involves first calculating mobility estimates using Equation (1) and then correlating these estimates with various location characteristics as a second-stage exercise.

Our approach instead follows the logic of intergenerational mobility theories (Cholli and Durlauf 2022) by incorporating both family and social variables into a unified model of child income. This model takes the form:

$$Y_i^c = \alpha(\boldsymbol{X}_i, \boldsymbol{S}_{-in}, n) + \beta(\boldsymbol{X}_i, \boldsymbol{S}_{-in}, n)Y_i^p + \varepsilon_i. \tag{2}$$

In this model, the mobility parameters $\alpha(\cdot)$ and $\beta(\cdot)$ are influenced by three factors: neighborhoods n as in Equation (1); individual family characteristics $\boldsymbol{X}_i$; and social characteristics $\boldsymbol{S}_{-in}$, which are derived from the characteristics of other families within social groups in the neighborhood. Conditioning on these characteristics, as well as control functions to account for selection on unobservables captured in the error term ε_i, is crucial for accounting for sorting that may drive heterogeneity across neighborhoods.

A key challenge of studying Equation (2) is its data demands: It requires linking not only the income of parents with their children, but also a host of family characteristics that form the basis of $\boldsymbol{X}_i$ and $\boldsymbol{S}_{-in}$. We overcome this challenge by leveraging rich administrative data from Denmark. This dataset provides detailed information on families for cohorts born between 1973 and 1983. It includes key characteristics including wealth, education, household structure, health, and crime. Using this rich data, we analyze how intergenerational mobility varies across 273 municipalities and nearly 2,000 parishes, which represent larger and smaller neighborhood units.

Our analysis proceeds in three parts. In the first part, we begin by estimating the workhorse mobility model of Equation (1) to provide baseline results of the heterogeneity of neighborhood mobility estimates. We conduct statistical hypothesis tests to assess how much neighborhood mobility varies. We then break down the observed distribution of mobility estimates into two parts: a "signal" component representing real differences in mobility, and a "noise" component caused by sampling error. This helps us distinguish between actual variations in mobility and random fluctuations in our data. Even among baseline estimates that exclude any controls, we find that at least 71% of municipalities and 88% of parishes have statistically indistinguishable AUM estimates from one another and that a significant fraction of variation is explained by noise.

We then demonstrate that the results of standard second-stage correlation exercises are sensitive to controlling for family or social characteristics in Equation (1). This dramatically affects interpretations of the mechanisms that drive neighborhood variation in mobility and motivates studying a generalized mobility model. For example, the share of non-Western immigrants in neighborhoods—the largest negative correlate of neighborhood AUM estimates from Equation (1)—is no longer associated with neighborhood AUM after controlling for the share of intact families among neighborhood social groups. The fragility of these exercises, coupled with the presence of large sampling error, motivates our ensuing analysis.

In the second part of the paper, we estimate linear specifications of Equation (2) that explicitly account for selection. We estimate Equation (2) using two alternative methods. The first employs a selection on observables assumption;[1] i.e., that controlling for family and social characteristics ($\boldsymbol{X}_i, \boldsymbol{S}_{-in}$) is sufficient to recover the neighborhood's location effects captured by the location-specific parameters. Though this method is simple to implement and analyze, families' selection into neighborhoods may depend on unobservable family-level preferences or the neighborhood's idiosyncratic features. We, therefore, also estimate Equation (2) accounting for selection on unobservables.[2] We construct control functions for neighborhood membership based on the residential histories of families that extend the method proposed by Dahl (2002).

These regressions allow us to assess the extent to which families' selection is the source of the neighborhood mobility differences relative to pure (residual) location effects. Specifically, we conduct a decomposition exercise to quantify the relative contributions of selection and location effects to variation in mobility found in our population and distinguish the role that sampling error may play in each of these components. Under either the selection on observables or unobservables approach, we find that selection explains the vast majority of the overall population distribution of mobility estimates. Meanwhile, the signal of municipality and parish location effects make up at most 4.3% and 2.9% of the overall variance, respectively. With respect to the between-neighborhood variance in mobility, observed family and social characteristics contribute at least 1.8–3.2 times more than location effects.

We also examine the spatial landscape of mobility from the baseline model (Equation 1) and our location effects from the generalized model (Equation 2). We find a surprising contrast between urban and rural areas. Urban areas that seemed to have low mobility according to the baseline model actually show high mobility when we look at the location effects from our generalized model. This reversal reflects the spatial autocorrelation of selection. We also find that residual location effects are extremely noisy: Hypothesis tests indicate that nearly all location effects of neighborhood-level AUMs are statistically indistinguishable from the AUM predicted by a counterpart population-level model. However, we find that there is still some meaningful variation in these location effects, even if it is small. This indicates that where a child grows up still influences their outcomes, but less than what traditional mobility analyses might suggest.

Third, we examine the association between the residual location effects obtained in the

1. See, e.g., Heckman and Navarro (2004).

2. See Heckman and Robb (1985)

second part of our analysis and location characteristics such as school quality, labor market structure, and rurality. This comparison helps us understand how model specification affects the interpretation of the sources of neighborhood heterogeneity. We find that many factors strongly linked to mobility in the baseline model, such as average home value, are no longer significantly associated with our location effects. This suggests that the associations in the baseline model are largely due to selection. In contrast, characteristics related to how rural or urban an area is are the strongest predictors of location AUM effects. This implies that in Denmark, the difference in mobility between locations is primarily explained by the different amenities available in urban versus rural areas. To account for estimation error, we implement empirical Bayes shrinkage methods using the neighborhood characteristics to attain more precise location effect estimates.

After completing these three steps, we have reached the limits that our linear specifications of the generalized mobility model can explain neighborhood heterogeneity in mobility, and any remaining unexplained differences in location effects is "irreducible." Our next task is to explore the properties of this irreducible heterogeneity. Here, we move in two directions. First, we examine the clustering properties of the irreducible location effects. We find evidence of the presence of two neighborhood "types": one with low location intercepts and steep location slopes in parental income, and one with the opposite features. Second, we argue that these irreducible location effects are suggestive of the need to more systematically explore nonlinearities in mobility dynamics than is conventional. Equation (2) posits that $\alpha(\cdot)$ and $\beta(\cdot)$ are general functions that may depend *nonlinearly* on family characteristics $\boldsymbol{X}_i$ and social characteristics $\boldsymbol{S}_{-in}$, as argued by Becker et al. (2018) and Durlauf and Seshadri (2018). Thus, the irreducible location effects derived from our linear specifications might reflect neighborhood averages of nonlinearities present in the mobility process (White 1980). We find evidence that parental income and mothers' labor force participation are strongly correlated with irreducible AUM, suggesting they nonlinearly affect child income.

In conclusion, this study offers a systematic examination of the heterogeneity in mobility across neighborhoods. While we find that the degree of variation in certain mobility measures, as defined in Equation (1), is comparable to that observed in other countries, we recognize that our findings may not be directly applicable to countries outside of Denmark. Nonetheless, we propose our methodological approach as a valuable model for investigating neighborhood mobility in other contexts. By developing a generalized mobility model and analyzing the structure of irreducible heterogeneity, we demonstrate how a more integrated approach combining economic theory and empirical analysis can shed light on the factors contributing to disparities in mobility across neighborhoods.

Relation to Existing Literature. Our analysis complements and extends a range of previous efforts to evaluate the sources of neighborhood mobility differences. While this paper addresses selection as an explanation for neighborhood heterogeneity in mobility, most of the empirical literature is purely descriptive and documents results similar to the first part of our paper. Chetty et al. (2014) and Chetty et al. (2018) estimate mobility in U.S. commuting zones and Census tracts using variants of Equation (1) and engage in second-stage correlation exercises on a range of location characteristics. They conclude that neighborhood levels of inequality, racial segregation, and family structure are strongly associated with mobility. This approach has been followed by a number of authors for different countries and continents, including Africa (Alesina et al. 2021), Australia (Deutscher and Mazumder 2020), Brazil (Britto et al. 2022), Canada (Corak 2020), Denmark (Eriksen and Munk 2020), France (Kenedi and Sirugue 2023), Germany (Dodin et al. 2024), Great Britain (Bell et al. 2022), Italy (Güell et al. 2018; Acciari et al. 2022), Sweden (Heidrich 2017), Switzerland (Chuard and Grassi 2020), and Turkey (Aydemir and Yazici 2019).

A recent literature initiated by Chetty and Hendren (2018a, 2018b) employs a "movers exposure" research design that compares outcomes of children who move to a given neighborhood at different points of childhood.[3] Under the assumption that selection effects do not vary by the age the child moves, this strategy identifies causal effects (among *movers only*) of the bundle of factors associated with neighborhoods. In particular, these identified effects may be driven by the composition of the self-selected population of permanent residents or place-based location characteristics. Deutscher (2020) and Laliberté (2021) employ additional quasi-experimental designs to identify the contribution of peer effects and school quality toward these neighborhood effects. In contrast to this body of work, the second part of our paper employs a unified modeling framework that allows us to decompose the competing roles of selection and location effects among the *full population*.[4]

With respect to the clustering of mobility estimates pursued in the final part of our paper, Connolly et al. (2019) and Corak (2020) are important predecessors of our analysis, which classify neighborhoods in the U.S. and Canada using a myriad of mobility statistics. Importantly, however, these authors cluster neighborhoods without accounting for selection, while we only cluster the irreducible location effects after exhausting our set of explanatory sources of neighborhood heterogeneity.

3. Papers employing this strategy include Alesina et al. (2021), Britto et al. (2022), Chetty et al. (2018), Chetty et al. (2020), Deutscher (2020), and Laliberté (2021).

4. Rothbaum (2016), Gallagher et al. (2019), Chetty et al. (2020), and Dodin et al. (2024) consider controlling for select family characteristics to address selection on observables into neighborhoods.

Organization. The paper progresses as follows. Section 2 describes our data and neighborhood sorting patterns that motivate our analysis. Section 3 estimates the baseline mobility model (Equation 1). Here, we assess the degree of heterogeneity of neighborhood mobility and the interpretation of second-stage correlation exercises. Section 4 estimates specifications of the generalized mobility model (Equation 2), the contributions of selection and location effects, and re-examines second-stage correlations with these location effects. Section 5 analyzes the structure of remaining irreducible heterogeneity. Section 6 concludes.

2 Data

Our study uses administrative data covering the full population of Denmark between 1980 and 2018. Unique individual identifiers allow us to link children with their parents and link families to a wide range of characteristics across time. In addition, individuals are linked with anonymized address identifiers each year, allowing us to measure neighborhood of residence. Appendices B.1 and B.2 describe the data and variables in detail.

2.1 Main Sample and Key Variables

We use the 1973–1983 birth cohorts to form our base sample. As mentioned above, children are linked with their parents (birth or adoptive). We exclude observations with missing parent identifiers, missing data before age 17, insufficient number of years or low levels of child or parent income, and children who are first- or second-generation immigrants.[5] Based on our neighborhood assignments (described below), we also exclude parishes with fewer than 25 families since they are prone to large estimation error in our analyses. This leaves us with over 560,000 parent-child dyads in our main sample.

Income. We link the sample to information on income. Most of our analyses focus on gross market income (earnings, self-employment income, and capital income), though we also consider income after transfers and after taxes in our robustness checks.

We use income registers between 1980–2018. Our preferred definition of child income is the log average income between ages 30–45 whenever income is available in the income registers; parent income is the log average sum of father and mother income when the child

5. Appendix Table B.1 summarizes how these restrictions affect our sample size. We restrict our sample to parents and children with mean income of at least $1,000. We also restrict our sample to native Danes since first-generation immigrants are not observed during their entire childhood while key variables (e.g., education) are subject to substantial measurement error for second-generation immigrants' parents; this reduces our sample by about 5%.

is ages 0–17 whenever available.[6] While log incomes are our primary focus given their ease of interpretation, we also consider other measures of socioeconomic status in our robustness checks. To mitigate the influence of outliers, we winsorize the upper 0.5% tails of the child and parent income distributions. All monetary values are expressed in 2010 U.S. dollars.

Neighborhoods. Using information on residential addresses, we link the sample to broader areas of residence. We consider two types of neighborhoods: parishes and municipalities. Parishes are administrative geographic units historically established by the Church of Denmark. There are over 2,100 parishes in Denmark, but due to our sample restrictions, our analysis considers 1,949 parishes. On average, a parish is home to about 2,500 residents (comparable to a small U.S. Census tract). Municipalities are larger governing bodies that manage local tax rates, school budgets, and social welfare programs. There are over 270 municipalities (before the 2007 municipal reform) that typically nest groups of parishes.[7] These units benefit from larger statistical power, providing an attractive way to examine if our results may be spuriously driven by small sample sizes of parishes. Children are assigned to the parish they lived in for the longest duration between ages 0–17; the municipalities containing the assigned parish determines municipality assignments.[8]

Family Characteristics. We also link children to a rich set of individual-level family background characteristics measured when the child is between ages 0–17, including education, household structure, marital status, labor market participation, assets, hospitalizations, and crime. Appendix B.2.3 provides a detailed description of all family characteristics used in our analysis.

Social Characteristics. Using these family characteristics and the neighborhood assignments, we construct social characteristics by taking averages of characteristics of *other* families in our sample that are members of the same social group within a parish or municipality.

6. This means child income for the 1973 birth cohort is defined as the log average between ages 30–45, while child income for the 1983 birth cohort is defined as the log average between ages 30–35. Similarly, parent income for the 1973 birth cohort is defined as the log average between ages 7–17, while parent income for the 1983 birth cohort is the log average between ages 0–17. We prefer using as many years of income available to mitigate issues on measurement errors arising from life-cycle bias (see, e.g., Mazumder 2005), though we also consider estimates of income over the same age ranges. Nevertheless, Appendix Figure B.1 demonstrates that there is roughly even distribution of birth cohorts across neighborhoods, suggesting life-cycle bias does not appear to be a major concern for our analysis.

7. Eriksen and Munk (2020) use 98 municipalities demarcated after the 2007 reform.

8. Since administrative records begin in 1980, our earliest birth cohort have data beginning only at age 7. Appendix Figure B.2 illustrates how families sort to the assigned parishes and municipalities over the child's life-cycle. It demonstrates that residential neighborhoods largely stabilize by age 7, giving confidence that there are no significant bias arises in our neighborhood assignment methodology across birth cohorts.

These "leave-out means" proxy for social influences and are customarily employed in the social interactions literature (Blume et al. 2011). Our specifications use social characteristics of two types of social groups. The first is school cohorts within a parish or municipality. Families are assigned to school cohorts based on the time their child enters 8th grade, which proxies their peers from schools and their own neighborhood. The second social group is small blocks, granular geographic units made up of roughly 150 households that proxy a family's local social network.[9]

Location Characteristics. Finally, we link neighborhoods to a set of location characteristics. These capture a neighborhood's overall socioeconomic status, labor market structure, demographics, and local public goods. Unlike social characteristics, many of the locational characteristics are statistics calculated on the *full population* of Danish adults living in the neighborhoods (rather than the main sample), typically averaged between 1980–2000. Characteristics include mean market income, the Gini index (in market income of adults between ages 18–60), mean housing value, home ownership rate, mean 9th grade test scores, share of commuters, urbanicity, share of farmers, and share of non-Western immigrants.[10]

2.2 Descriptive Patterns of Neighborhood Stratification

To motivate our analysis, we begin with descriptive patterns that illustrate how neighborhoods embody inequalities in family, social, and location characteristics. Appendix Tables C.1–C.3 provide summary statistics of all family, social, and location characteristics.

Figure 1 summarizes the key issues at hand. Figure 1(a) plots parish-level averages of standardized measures of child income, parental income, mother education level, intact family structure, hospitalizations, and crime by the parish's rank in each characteristic. For all characteristics, the lowest and highest ranked neighborhoods tend to differ by 1–2 standard deviations, showcasing how Danish neighborhoods possess significant differences in their composition of families due to sorting. Focusing on parental income, we find the mean parental income of the middle 90% of parishes lies between \$55,600 and \$81,800, corresponding to the middle 44% of the population income distribution.[11] For many other

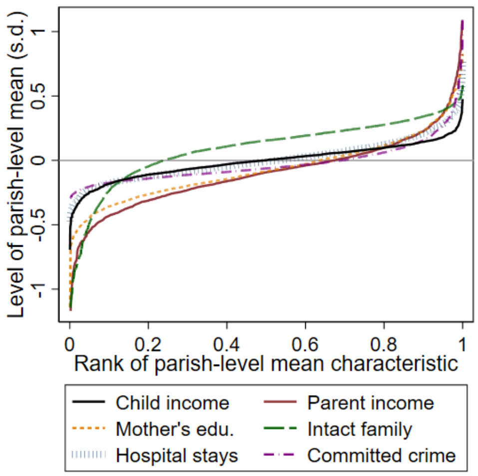
9. See Damm and Schultz-Nielsen (2008) for details on the construction of these small block units and see Appendix B.2.4 for more details on social characteristics. We have checked that our results are robust to using social characteristics that use only one of these social groups or alternative social group definitions such as families residing in large blocks or, in the case when the neighborhood unit is municipalities, parishes.

10. Ninth grade test score data is available only starting in 2002. Figure B.3 demonstrates that mean neighborhood test scores is strongly correlated with measures of neighborhood's socioeconomic status over time, indicating it provides a strong signal of mean test scores experienced by our study's cohorts.

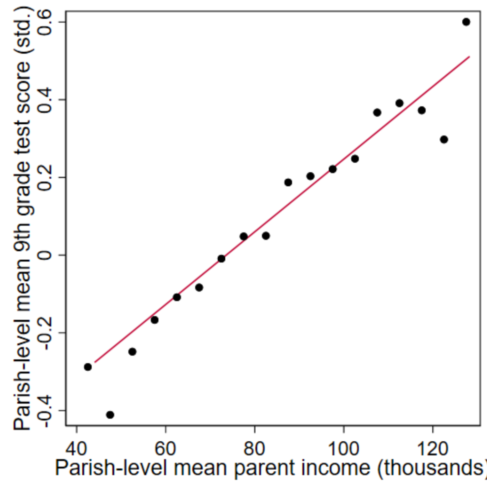
11. See Appendix Figures C.1 and C.2 for plots of these distributions.

Figure 1: Inequality in family and location characteristics

(a) Inequality in family characteristics across parishes



(b) Association between mean 9th grade test scores and parental income across parishes



Notes: Panel (a) plots within-parish mean family attributes, ranked from smallest to largest. Family characteristics are standardized over the pooled sample. Panel (b) presents the association between parish-level mean parental income and parish-level mean (nationally-standardized) 9th grade test score. The scatter plot are means within \$5,000 mean parish income bins; these means and the linear fit (red line) are weighted by parish sample sizes. See Appendix Figure A.1 for corresponding plots for municipalities.

characteristics, Theil index decompositions reveal that parishes explain meaningful levels of inequality across families; see Appendix Table C.4.

Figure 1(b) highlights how another possible driver of neighborhood heterogeneity in mobility—location characteristics such as mean standardized test scores—are confounded by selection. The plot documents a strong positive association between mean parental income and 9th grade test scores, revealing how neighborhood sorting may explain the relationship between common measures of local public goods and mobility. These descriptive statistics highlight the importance of disentangling selection from sorting of families and location effects driven by local public goods like schools for understanding differences in mobility across neighborhoods.

3 Baseline Results from the Workhorse Mobility Model

For the first stage of our analysis, we estimate Equation (1) to provide a set of baseline results that we can use as a benchmark against later specifications that account for selection on observables and unobservables. These results also provide an opportunity to compare the magnitude of heterogeneity of neighborhood mobility estimates with what has been documented in other countries like the U.S.

Throughout our paper, we focus on two key neighborhood mobility measures based on parent and child incomes expressed in logs:[12] the *intergenerational elasticity* of income (IGE), corresponding to the neighborhood-specific slope coefficient β_n, and the *absolute upward mobility* (AUM) $\bar{y}^c_{qn} \equiv \alpha_n + \beta_n y^p_q$, which is the predicted log income for a child that grows up in neighborhood n conditional on setting parental income to $Y^p_i = y^p_q$, the qth percentile of the national parent income distribution. In much of our analysis, we use $q = 25$, since the bottom income quartile is a common benchmark of socioeconomic disadvantage. We sometimes exponentiate the AUM to express our results of predicted child income as levels instead of logs.[13]

3.1 Variation in Baseline Estimates of Neighborhood Mobility

Raw Mobility Estimates. We first focus on descriptive patterns in the raw estimates from Equation (1). Table 1 reports summary statistics of neighborhood mobility point estimates.[14] The modal parish and municipality has an IGE estimate of around 0.32 and an AUM estimate of $33,400, but there is considerable spread across neighborhoods. Among municipalities (reported in Panel A), the standard deviation of municipality-level IGE estimates is 0.06. A one standard deviation increase in a child's municipality-level AUM is associated with an increase of about $1,160 per year in predicted income during adulthood. Parishes exhibit even greater variation, with a standard deviation change associated with about a 0.18 change in the IGE and a $2,860 change in the AUM.

To better quantify the localized nature of variation in mobility estimates, Appendix Table A.1 decomposes the variance of parish-level mobility estimates within and between municipalities. The variance of parish estimates *within municipalities* contributes around 80% of the total variance, highlighting how granular neighborhood units appear to have large consequences in a child's future outcomes. For example, among parishes within Aarhus—the second largest municipality in Denmark—AUM estimates range between $26,100 (corresponding to the 1st percentile of the parish AUM distribution) and $37,600 (corresponding

12. We focus on the log–log regression model given its popularity and ease of interpretation. Appendix D provides results on alternative mobility statistics, such as rank–rank regression estimates and transition probabilities.

13. Both neighborhood-level statistics capture distinct dimensions of intergenerational mobility. The neighborhood-level IGE captures a notion of *relative* income mobility: what is the associated change in child income due to a 1% increase in parental income in neighborhood n? Meanwhile, the neighborhood-level AUM captures a notion of *absolute* income mobility: what is the child's predicted income in adulthood in neighborhood n, conditional on their parents lying at the bottom income quartile?

14. Eriksen and Munk (2020) find similar patterns in their intergenerational mobility estimates across the 98 municipalities defined after the 2007 municipality reform.

Table 1: Summary Statistics of Neighborhood-level Mobility Estimates

	N	Mean	S.D.	Min.	Percentiles 10th	25th	50th	75th	90th	Max.
A. Municipality-level mobility measures										
$\hat{\beta}_n$	273	0.328	0.062	0.139	0.248	0.292	0.332	0.369	0.402	0.507
$\hat{\bar{y}}_{25n}$ (thousands)	273	33.4	1.16	30.3	32	32.6	33.4	34.2	35.1	36.5
Sample size	273	2,055	2,689	313	678	930	1,276	2,193	4,047	25,418
B. Parish-level mobility measures										
$\hat{\beta}_n$	1,949	0.316	0.179	−0.727	0.114	0.218	0.315	0.406	0.522	1.21
$\hat{\bar{y}}_{25n}$ (thousands)	1,949	33.4	2.86	23.3	30.1	31.7	33.3	35.2	37	49.6
Sample size	1,949	288	344	25	43	74	155	363	720	3,117

Notes: This table reports summary statistics of estimates of the IGE ($\hat{\beta}_n$) and AUM ($\hat{\bar{y}}^c_{25n}$) from Equation (1) and the number of observations in municipalities and parishes from our main sample. All AUM estimates $\hat{\bar{y}}^c_{25n}$ are exponentiated and expressed in thousands of 2010 U.S. dollars.

to the 93rd percentile), suggesting that neighborhood upbringing may matter at a highly granular level. In fact, Appendix Figure A.2 plots mobility estimates among large and small blocks, which are composed of roughly 600 and 150 households, and finds even larger variance in point estimates.[15] To visualize this spatial variation, Appendix Figure A.3 illustrates heat maps of the municipality- and parish-level IGE and AUM point estimates.

These descriptive patterns of differences in mobility across neighborhoods are corroborated by a series of robustness checks. We find qualitatively similar results of mobility estimates based on rank-rank regressions (Chetty et al. 2014), level–level regressions, transition probabilities, and Poisson–transformed regressions (Mitnik and Grusky 2020); see Appendix D. We also find qualitatively similar variation in educational mobility (where Y^c_{in}, Y^p_{in} are defined as years of schooling). While using disposable income (after taxes and transfers) decreases heterogeneity in income mobility across neighborhoods, the results remain highly correlated with estimates using market income.[16]

Clearly, neighborhood-level variation in mobility estimates is a prevalent feature in Denmark, just as in developed countries with less prominent welfare states. In fact, the heterogeneity of certain income mobility estimates across Danish parishes falls within a comparable order of magnitude as the U.S. Census tracts reported by the Opportunity Atlas (Chetty et al. 2018). Appendix Figure G.1 demonstrate that distributions of rank-based neighbor-

15. We exclude blocks with fewer than 25 families in our main sample.

16. See Appendix Table D.1, Panel B.

hood slope estimates (similar to the IGE) between both countries are quite similar, while the distribution of AUM estimates in Denmark is more compressed.

Accounting for Sampling Error. The large observed heterogeneity even *within* municipalities might suggest that the specific location that a child is raised in is critical in determining their future socioeconomic status. However, sampling error may inflate the variation of neighborhood mobility estimates and mask the actual degree of heterogeneity in the model's true neighborhood mobility parameters. This is especially a concern when considering more granular neighborhood units like parishes, which are prone to having a small numbers of families in our sample.

In Appendix E, we critically assess how sampling error drives this heterogeneity through three types of statistical exercises. The first two exercises involve joint and multiple hypothesis test procedures (MHTP) that establish the extent that *individual neighborhoods* vary in mobility; to the best of our knowledge, these have not been pursued in the neighborhood mobility literature. First, Appendix Table E.1, Panel A reports that at least 71% of municipalities and 88% of parishes have statistically indistinguishable AUM estimates from one another.[17] This is consistent with Andrews et al. (2024) and Mogstad et al. (2024), who document substantial uncertainty around the rankings of AUM estimates among U.S. neighborhoods. Second, Panel B reports shares of neighborhood mobility estimates that are statistically indistinguishable from mobility estimates from the population-level mobility regression ($Y_i^c = \alpha + \beta Y_i^p + \varepsilon_i$) based on the Benjamini and Hochberg (1995) step-up MHTP, which controls for the probability of "false discoveries" or incorrect rejections of the null. Remarkably, only about 5% of municipalities and 3% of parishes possess AUM estimates that are statistically different from the pooled AUM estimate ($\hat{\bar{y}}_{25}^c \equiv \hat{\alpha} + \hat{\beta} y_{25}^p$).

The third exercise aims to assess the magnitude of true parameter heterogeneity among the *distribution of neighborhoods*. Following practices pursued by some of the neighborhood mobility literature (e.g., Chetty and Hendren 2018b, Chetty et al. 2018), we conduct standard signal–noise decompositions to quantify the fraction of variance in neighborhood mobility estimates explained by sampling error (e.g., $\frac{\mathbb{E}[\hat{\sigma}^2_{\hat{\beta}_n}]}{\mathrm{Var}(\hat{\beta}_n)} \times 100\%$, where $\hat{\sigma}_{\hat{\beta}_n}$ is the standard error of $\hat{\beta}_n$).[18] Appendix Table E.4 finds that about 38% of the unweighted variance of municipality AUM estimates and 72% of the unweighted variance of parish AUM estimates are driven by noise; weighing neighborhoods by sample sizes reduces the role of noise, but it clearly remains a significant driver of neighborhood heterogeneity. Relative to our hypothesis tests, the signal-noise decomposition results provide a somewhat more optimistic outlook of the

17. This result is based on an iterative series of joint Wald tests; see Appendix E for details.

18. This is the complement of what is sometimes called the "reliability score."

degree that neighborhood mobility estimates possess signal.

In sum, these exercises demonstrate how estimation error—even in the baseline mobility model—generates significant spurious heterogeneity in neighborhood mobility estimates. They highlight the importance of accounting for noise when we turn to estimating more sophisticated models of mobility and decomposing its underlying sources. They also suggest the promise in identifying properties of the *distribution* of neighborhood mobility rather than *individual* neighborhood mobility, such as through second-stage correlation exercises.

3.2 Second-Stage Correlations and their Interpretation

With the neighborhood-level mobility estimates from Equation (1) at hand, a common practice in the empirical mobility literature is to correlate these estimates with various location characteristics $\boldsymbol{Z}_n$. These second-stage correlation exercises aim to provide suggestive evidence of the underlying drivers of heterogeneity in mobility across neighborhoods. Typically, second-stage correlation exercises are based on bivariate regressions focusing on one element of $\boldsymbol{Z}_n$ at a time; in the case of the neighborhood IGE,

$$\hat{\beta}_n = \delta_0 + \delta_1 Z_n + u_n. \tag{3}$$

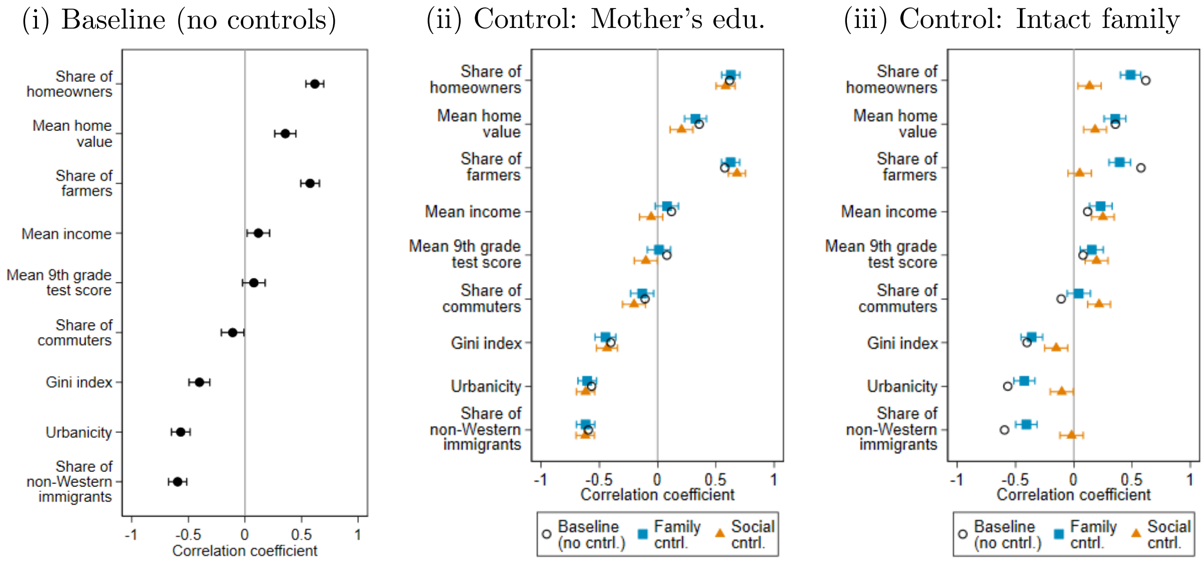
Scaling δ_1 by the ratio of the standard deviations of $\hat{\beta}_n$ and Z_n delivers the correlation coefficient.

Results. Following the literature, the left panels of Figure 2 report second-stage correlations between neighborhood AUM estimates and a multitude of neighborhood characteristics for both municipalities and parishes; Appendix Figure A.4 reports results for the IGE. We find that neighborhoods with higher mobility are associated with higher mean income, home value, and homeownership rates; higher 9th grade test scores; higher shares of farmers and lower shares of non-Western immigrants; and lower income inequality as measured by the Gini index, which is consistent with an intranational "Great Gatsby Curve" (Krueger 2012; Corak 2013; Durlauf and Seshadri 2018). These findings are broadly consistent with patterns found in other countries.

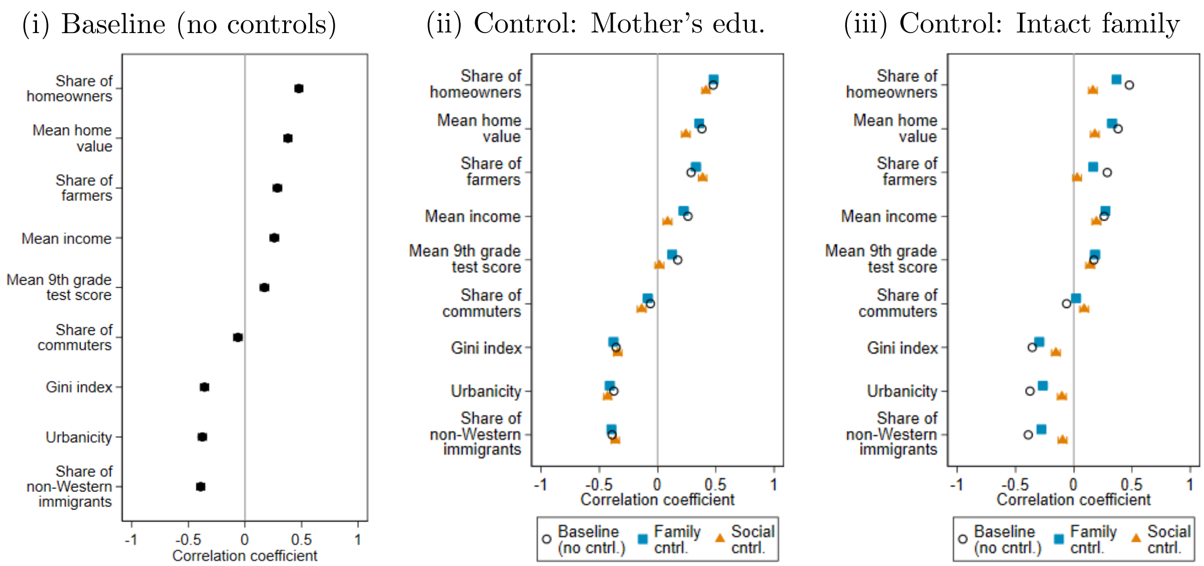
Interpretation. Given that both the mobility estimates and location characteristics used in calculating the correlations presented above are measured at the neighborhood level, one may be tempted to interpret their correlations as reflections of strictly "place-based" features of mobility. However, most location characteristics are arguably correlated with their self-selected populations, or are themselves defined as neighborhood mean family characteristics

Figure 2: Second-stage correlations between location characteristics and neighborhood AUM estimates, without vs. with select controls

(a) Municipalities



(b) Parishes



Notes: This figure plots correlation coefficients between neighborhood-level AUM estimates and location characteristics. Left panels plot correlations using baseline neighborhood AUM estimates $\bar{y}^c_{25n}$ (from Equation (3)). Middle and right panels show correlations using neighborhood AUM estimates after controlling for mother's education level and intact family as in Equation (4). Square markers measure controls at the individual family level (i.e., components of $\boldsymbol{X}_i$); triangular markers represent measure controls at the social level, among other families of children belonging to the same birth cohort in the neighborhood (i.e., components of $\boldsymbol{S}_{-in}$). Horizontal capped lines are 90% confidence intervals.

$\boldsymbol{X}_i$. Other characteristics commonly studied include segregation, institutional quality, and local labor markets that may not be directly constructed from $\boldsymbol{X}_i$ but could be potentially affected by $\boldsymbol{X}_i$. This raises serious concerns over the interpretation of these correlation exercises as descriptions of mobility aspects that are place-specific.

To put it differently, in light of Figure 1, it is clear that selection may directly affect the neighborhood mobility estimates from Equation (1) and their association with location characteristics. To help fix ideas, consider the simple case where the true data generating process of neighborhood mobility includes the mother's educational attainment Edu_i,[19]

$$Y_i^c = \tilde{\alpha}_n + \tilde{\beta}_n Y_i^p + \tilde{\gamma} Edu_i + \epsilon_i, \tag{4}$$

where $\tilde{\gamma} \neq 0$. Notice that, in this model, mother's education has a uniform effect on child income irrespective of the family's neighborhood. This permits us to examine an "individual family effect" (γ) that is separate from the "location effect," which consists of both the direct effect on the child α_n and the effect mediated through parent income β_n.

In this case, the identified neighborhood IGE from Equation (1) is

$$\beta_n = \tilde{\beta}_n + \theta_n^{Edu},$$

where $\theta_n^{Edu} \equiv \tilde{\gamma} \frac{\mathrm{Cov}\left(Y_i^p, Edu_i \mid n(i)=n\right)}{\mathrm{Var}\left(Y_i^p \mid n(i)=n\right)}$ is the omitted variable bias (OVB) that represents the within-neighborhood gradient of parent income and mother's education. Thus, the identified neighborhood IGE includes both the pure "location effect" as well as the "individual family effect" captured through the parent income–mother's education gradient within neighborhoods. This implies that the second-stage slope coefficient recovered from Equation (3) will similarly comprise both a location and an individual family effect. Similar OVB arguments can be applied to α_n to decompose the correlations with neighborhood AUM $\bar{y}_{25n}^c$ into neighborhood and individual family effect components as well.

Though the neighborhood mobility measures and their concomitant correlations identified by Equations (1) and (3) provide interesting descriptive patterns, it is clear from this example that they are difficult to interpret. Despite the fact that second-stage correlation exercises involve variables measured at the neighborhood level—even place-based characteristics that may generate a pure location effect—the correlations themselves might express effects driven by selection, be it through individual family effects (as captured through θ_n^{Edu} in the example) or social effects from the neighborhood's local population. This can lead to misleading conclusions about the underlying sources of neighborhood heterogeneity. Our

19. Arguments for richer models that include other family characteristics and social characteristics $\boldsymbol{S}_{-in}$ can be extend naturally.

simple example highlights how the degree that OVB arising from selection is associated with location characteristics might affect the magnitude, sign, and interpretation of second-stage correlation exercises.

In practice, we find such bias can be economically and statistically significant. The middle and right panels of Figure 2 illustrate correlation coefficients between location characteristics and AUM estimates following the example in Equation (4). As controls, we use different family and social characteristics that reflect selection: mother's education level and intact family structure. When measured at the family or social level, these characteristics proxy for different theories of intergenerational mobility—family channels of human capital investment or social influences through role model effects—that are distinct from theories of strictly location effects. We find that controlling for mother's education at family or social level (middle panels) has little impact but controlling for intact family structure at the social level (right panels) qualitatively changes associations relative to baseline. For example, the neighborhood's share of farmers and non-Western immigrants—which are respectively among the strongest positive and negative correlates of the baseline AUM—are no longer correlated with the AUM upon including this control. This reveals that interpreting correlations with baseline AUM estimates from the lens of place-based theories, such as the neighborhood's occupational structure or degree of ethnic diversity, may be misplaced. These results show how baseline correlations can be driven by associations between parent income and omitted variables that reflect selection.

This analysis highlights that the relative contribution of selection toward baseline mobility estimates can be large. Changes in correlation patterns from controlling for a single omitted family or social characteristic suggest that second-stage correlation exercises with baseline mobility estimates from Equation (1) are rather fragile. This implies that second-stage correlation exercises typically pursued in the literature should not be interpreted as uncovering mechanisms of how location characteristics explain mobility differences across neighborhoods. Controlling for selection using a rich set of family and social characteristics will be critical to retrieve correlations that appropriately reflect the actual locational drivers of neighborhood mobility.

4 Results from a Generalized Mobility Model

The baseline neighborhood mobility model, Equation (1), does not explicitly control for sorting. Rather, the average effect of selection—from both observable and unobservable characteristics—on child income is subsumed by the neighborhood index $n \in \mathcal{N}$ in the location parameters (α_n, β_n). As we have just shown, this poses issues in interpreting the underlying mechanisms that generate neighborhood mobility. This raises the question: How much does selection arising from family and social characteristics explain variation in neighborhood mobility relative to location effects? How much residual neighborhood heterogeneity in mobility exists after explicitly controlling for selection?

To answer these questions, we generalize the workhorse neighborhood mobility model by estimating a linear specification of Equation (2),

$$Y_i^c = \underbrace{\alpha_n + \boldsymbol{\alpha_x X}_i + \boldsymbol{\alpha_s S}_{-in}}_{\alpha(\boldsymbol{X}_i, \boldsymbol{S}_{-in}, n)} + \underbrace{(\beta_n + \boldsymbol{\beta_x X}_i + \boldsymbol{\beta_s S}_{-in})}_{\beta(\boldsymbol{X}_i, \boldsymbol{S}_{-in}, n)} \cdot Y_i^p + \varepsilon_i, \tag{5}$$

where $\boldsymbol{X}_i$ is individual family characteristics and $\boldsymbol{S}_{-in}$ is social-level characteristics (see Section 2.1).[20] In this model, the residual neighborhood parameters (α_n, β_n) represent the location effect that affect child income above and beyond our controls for selection. We focus on linear specifications due to the curse of dimensionality, but revisit the possibility of nonlinearities later in the paper.

Note that the family and social characteristics are specified to have common effects across all neighborhoods (i.e., the parameters $(\boldsymbol{\alpha_x}, \boldsymbol{\alpha_s}, \boldsymbol{\beta_x}, \boldsymbol{\beta_s})$ are not indexed by n) so that location effects are assumed to be additively separable from family and social characteristics. If interactions between $(\boldsymbol{X}_i, \boldsymbol{S}_{-in})$ and index n exist, they will be captured by ε_i. Though this is a strong assumption, we pursue this specification on conceptual and practical grounds. Conceptually, it allows us to bound the magnitude of residual neighborhood heterogeneity after controlling for selection. Practically, there are insufficient degrees of freedom to estimate a model where family and social characteristics have location-specific effects, and this specification permits a clean decomposition of neighborhood mobility along the dimensions of selection and location effects.

We begin by discussing our methodologies for estimating Equation (5). Next, we report estimates of the effects of selection on neighborhood-level mobility and re-assess the magnitude of residual neighborhood heterogeneity. We then conduct a decomposition exercise that disaggregates variation in mobility estimates along its constituent components of selection effects, location effects, and sampling error. Finally, we re-examine second-stage correlation exercises using our location effect estimates.

20. In order to compare the effects between components of $\boldsymbol{X}_i$ and $\boldsymbol{S}_{-in}$ on a common scale, all continuous characteristics are standardized so that they are expressed as standard deviation units. All dummy indicators are demeaned. Demeaning helps retain the centering of the neighborhood-level intercepts α_n.

4.1 Empirical Implementation

We work with two alternative assumptions for analyzing Equation (5): selection on observables and selection on unobservables.

Assumption O. *(Selection on observables.)* $\mathbb{E}\left[\varepsilon_i \mid \boldsymbol{X}_i, \boldsymbol{S}_{-in}, Y_i^p, n(i)=n\right] = 0.$

Assumption O states that any remaining unobservables captured by ε_i do not bias our estimates of selection and location effects. Under this assumption, we can directly identify the parameters of Equation (5) via ordinary least squares. Though convenient for empirical implementation, this is a strong assumption. It posits there are no interactions between $(\boldsymbol{X}_i, \boldsymbol{S}_{-in})$ and index n, and that there are no other family or social characteristics that influence child outcomes that are not fully captured by $(\boldsymbol{X}_i, \boldsymbol{S}_{-in})$.

Assumption U. *(Selection on unobservables.)* $\mathbb{E}\left[\varepsilon_i \mid \boldsymbol{X}_i, \boldsymbol{S}_{-in}, Y_i^p, n(i)=n\right] = K_n(\{P_{ij}\}_{j\in\mathcal{N}})$ *where* $P_{ij} \equiv \Pr(n(i)=j \mid \boldsymbol{X}_i, \boldsymbol{S}_{-ij}, Y_i^p, Z_{ij})$ *and* Z_{ij} *is a neighborhood* j*-specific instrument.*

In contrast to Assumption O, Assumption U allows unobserved factors to affect child income. It expresses the bias term $\mathbb{E}[\varepsilon_i \mid \boldsymbol{X}_i, \boldsymbol{S}_{-in}, Y_i^p, n(i)=n]$ that arises from selection on unobservables through a control function that depends on the probabilities of selecting each possible neighborhood, $K_n(\{P_{ij}\}_{j\in\mathcal{N}})$. By including a control function in our regression (5), we can properly identify the parameters of interest free of bias.[21]

The key challenge of Assumption U is its empirical implementation. We employ a semi-parametric control function approach that involves two steps: estimating the neighborhood choice probabilities $\{P_{ij}\}_{j\in\mathcal{N}}$ and specifying the function $K_n(\{P_{ij}\}_{j\in\mathcal{N}})$ in a computationally tractable way.

For the first step, we consider a multinomial discrete choice problem where parents who originate from neighborhood n_0 are selecting the neighborhood n that their child will reside for the majority of their childhood. For our neighborhood-specific instruments $\{Z_{ij}\}_{j\in\mathcal{N}}$, we use the geographic distance between n_0 and each potential neighborhood choice n. Note that each P_{in} depends generally on $\boldsymbol{X}_i, \boldsymbol{S}_{-in}, Y_i^p, Z_{in}$. To make this computationally tractable, we make two simplifications. First, we discretize a subset of family characteristics $\boldsymbol{X}_i$ and parent income Y_i^p to form 360 different "family types" indexed by $t \in \mathcal{T}$.[22] Second, we

21. Strictly speaking, this allows us to identify all parameters except the neighborhood intercept parameters α_n. These intercepts can be recovered through identification at infinity arguments: $\lim_{P_{in}\to 1} K_n(\{P_{ij}\}_{j\in\mathcal{N}}) = 0$. We discuss our approach for identifying α_n below.

22. Specifically, we discretize mother's age into terciles; mother's highest educational attainment based on less than high school, high school, or college plus; and parent income into quintiles and consider these variables along with indicators of whether the child resided in an intact household, either parent committed a crime, and either parent experienced inpatient hospitalization. The Cartesian product of these family characteristics results in $|\mathcal{T}| = 360$.

abstract away from social characteristics $\boldsymbol{S}_{-in}$.[23]

Together, these two simplifying assumptions dramatically reduce the dimensionality of the probabilities to $P_{in} = \Pr(n(i) = n \mid \boldsymbol{X}_i, \boldsymbol{S}_{-in}, Y_i^p, Z_{in}) \approx \Pr(n(i) = n \mid t(i), Z_{in})$. Assuming that neighborhood-specific preference shocks are drawn independently from a type-I extreme value distribution conditional on family type t, we compute the neighborhood choice probabilities via a per-type logit: letting $\mathcal{N}_t$ be the set of neighborhoods chosen by family type t, we have that $P_{in} = \frac{\exp\{\theta_{nt}+\xi_t Z_{in}\}}{\sum_{j\in\mathcal{N}_t}\exp\{\theta_{jt}+\xi_t Z_{ij}\}}$, where θ_{jt} are neighborhood-by-type-specific mean utilities of each neighborhood alternative j and ξ_t is type-specific disutility from distance to alternatives.

The second step involves specifying the control function $K_n(\cdot)$. Due to the curse of dimensionality, it is practically impossible to allow $K_n(\cdot)$ to nonparametrically depend on all the neighborhood choice probabilities $\{P_{ij}\}_{j\in\mathcal{N}}$. Different assumptions are invoked to reduce its dimensionality (Bourguignon et al. 2007). One recent approach employs the linearity assumption of Dubin and McFadden (1984), which states that the control function linearly depends on particular functions of each P_{ij}, and further assumes they are homogeneous across neighborhoods (i.e., $K_n(\cdot) = K(\cdot)$).[24] However, in our context, this assumes away interactions between $(\boldsymbol{X}_i, \boldsymbol{S}_{-in})$ and the neighborhood index n. An alternative approach applies the index sufficiency assumption of Dahl (2002), which states that $K_n(\cdot)$ depends on only a subset of $\{P_{ij}\}_{j\in\mathcal{N}}$ (e.g., the first few highest order statistics). This approach allows $K_n(\cdot)$ to depend on the subset of probabilities nonlinearly and to vary across neighborhoods.

We adopt a middle ground approach. We reduce the dimensionality of $K_n(\cdot)$ by exploiting the nested nature of Denmark's neighborhoods. Specifically, we assume that the control function depends on summary statistics of conditional-on-nest neighborhood and marginal nest choice probabilities.

Recall that parishes are nested within municipalities. We further classify municipalities into eight mutually exclusive and exhaustive "supernests" we call regions. A given parish n is nested (i.e., contained) in a municipality that includes other parishes as well. We denote the set of parishes contained in this municipality by $\mathcal{N}^1(n)$. This municipality, which contains parish n, is itself a member of a given region, which includes other municipalities as well. We let $\mathcal{N}^2(n)$ denote the collection of municipality nests that represents the region supernest containing parish n. Similarly, when we perform the analysis at the municipality level instead (i.e., when n indexes municipalities), then $\mathcal{N}^1(n)$ denotes the collection of municipalities lying

23. Introducing these variables requires satisfying equilibrium conditions that arise from a social interactions game where each individual family's neighborhood choice depends on the potential choices that other families make. Estimating such a model with such a large set of neighborhoods is challenging and beyond the scope of this paper.

24. Abdulkadiroğlu et al. (2020) and Otero et al. (2021) employ this strategy with an ordered logit model.

within the region nest containing municipality n. Let $\mathcal{N}^\ell(n)$ denote the ℓth-order nesting of neighborhood n, and define $\mathcal{N}^0(n) \equiv n$ and $\mathcal{N}^{L+1}(n) \equiv \mathcal{N}$. Then, there are $L = 1$ levels of nesting when we define municipalities as our neighborhood unit, and $L = 2$ levels of nesting when we study parishes.

Define the marginal nest and supernest probabilities as $P_{i\mathcal{N}^1(n)} \equiv \sum_{j \in \mathcal{N}^1(n)} P_{ij}$ and $P_{i\mathcal{N}^2(n)} \equiv \sum_{\mathcal{N}^1 \in \mathcal{N}^2(n)} P_{i\mathcal{N}^1}$. Further define the conditional-on-nest neighborhood choice and conditional-on-supernest nest choice probabilities as $P_{in\,|\,\mathcal{N}^1(n)} \equiv \frac{P_{in}}{P_{i\mathcal{N}^1(n)}}$ and $P_{i\mathcal{N}^1\,|\,\mathcal{N}^2(n)} \equiv \frac{P_{i\mathcal{N}^1}}{P_{i\mathcal{N}^2(n)}}$. Note that $P_{i\mathcal{N}^L(n)\,|\,\mathcal{N}^{L+1}(n)} = P_{i\mathcal{N}^L(n)}$. With these probabilities at hand, we specify the control function as

$$K_n(\{P_{ij}\}_{j \in \mathcal{N}}) = \underbrace{\sum_{\ell=1}^{L+1} \lambda_{\ell n} \cdot [1 - P_{i\mathcal{N}^{\ell-1}(n)\,|\,\mathcal{N}^\ell(n)}]}_{\substack{\text{Conditional choice probabilities} \\ \text{of chosen neighborhood}}} + \underbrace{\sum_{\ell=1}^{L+1} g_\ell(\{P_{i\mathcal{N}^{\ell-1}\,|\,\mathcal{N}^\ell(n)}\}_{\mathcal{N}^{\ell-1} \in \mathcal{N}^\ell(n)})\}}_{\text{All conditional choice probabilities}}. \qquad (6)$$

The first term allows the conditional choice probabilities of the chosen neighborhood n and its higher-order nests $\mathcal{N}^\ell(n)$ to have neighborhood-specific effects $\lambda_{\ell n}$, permitting selection on unobservables to vary by the index n. The second term's $g_\ell(\cdot)$ functions map each set of conditional choice probabilities to various summary statistics of the distributions of these probabilities, including variance, skewness, and kurtosis; 25th, 50th, and 75th quantiles; and second- and third-order statistics. Each of these summary statistics are assumed to have common effects across neighborhoods. This specification is more parsimonious yet permits the control function to indirectly depend on all of the neighborhood choice probabilities via the conditional choice probabilities.

Notice that, in the first term, $\lim_{P_{in} \to 1} [1 - P_{i\mathcal{N}^{\ell-1}(n)\,|\,\mathcal{N}^\ell(n)}] = 0$ for each ℓ. We also recenter each summary statistic in the second term so that $\lim_{P_{in} \to 1} g_\ell(\{P_{i\mathcal{N}^{\ell-1}\,|\,\mathcal{N}^\ell(n)}\}_{\mathcal{N}^{\ell-1} \in \mathcal{N}^\ell(n)}) = 0$.[25] Specifying $K_n(\cdot)$ in this fashion means that the control function possesses no intercept, obviating practical issues associated with recovering the neighborhood-level intercepts α_n via "identification at infinity" approaches (Chamberlain 1986, Heckman 1990).

For inference, we re-estimate the two-step procedure for estimating Equation (5) for 200 weighted bootstraps à la Rubin (1981). Standard errors are estimated with the standard deviations of estimates across bootstrap draws.

25. For example, we recenter the variance of $\{P_{ij\,|\,\mathcal{N}^1(n)}\}_{j \in \mathcal{N}^1(n)}$ of each family i with the value of the variance that sets $P_{in} = 1$ and $P_{ij} = 0$, $\forall j \in \mathcal{N}^1(n)$ s.t. $j \neq n$.

4.2 Estimates from the Generalized Mobility Model

Having estimated Equation (5), we compute three types of mobility statistics. To simplify notation, let $\boldsymbol{W}_{in} \equiv (\boldsymbol{X}_i, \boldsymbol{S}_{-in})$, $(\boldsymbol{\alpha_w}, \boldsymbol{\beta_w}) \equiv ((\boldsymbol{\alpha_x}, \boldsymbol{\alpha_s}), (\boldsymbol{\beta_x}, \boldsymbol{\beta_s}))$, and $K_{in} \equiv K_n(\{P_{ij}\}_{j\in\mathcal{N}})$.[26]

The first type is the *individual-level* IGE and AUM—that is $\beta(\boldsymbol{W}_{in}, n)$ and $\bar{y}^c_{i,25n} \equiv \alpha(\boldsymbol{W}_{in}, n) + \beta(\boldsymbol{W}_{in}, n)y^p_{25}$ (under Assumption O) or $\bar{y}^c_{i,25n} \equiv \alpha(\boldsymbol{W}_{in}, n) + \beta(\boldsymbol{W}_{in}, n)y^p_{25} + K_{in}$ (under Assumption U). The individual-level mobility parameters depend on both the location index n and the family and social characteristics $\boldsymbol{W}_{in}$, thereby capturing heterogeneity of intergenerational mobility experienced by individual children even within a given neighborhood n.

The second type of mobility statistic is the *mean neighborhood-level* IGE and AUM, which are defined as $\mathbb{E}[\beta(\boldsymbol{W}_{in}, n) \,|\, n(i) = n]$ and $\mathbb{E}[\bar{y}^c_{i,25n} \,|\, n(i) = n]$. This describes average mobility by neighborhood, analogous to the baseline estimates from Equation (1).

The third type of statistic is given by the two components of the individual-level mobility parameters. The first is the *selection effect*, which itself is decomposed into two terms: the *observable selection effect* of family and social characteristics ($\boldsymbol{\beta_w}\boldsymbol{W}_{in}$ and $\boldsymbol{\alpha_w}\boldsymbol{W}_{in} + \boldsymbol{\beta_w}\boldsymbol{W}_{in}y^p_{25}$) and the *unobservable selection effect* captured by the control function. The second copmonent is the *location effect*, i.e., the residual effect of neighborhoods after accounting for selection (β_n and $\bar{y}^c_{25n} \equiv \alpha_n + \beta_n y^p_{25}$). The key difference between the location effects β_n and $\bar{y}^c_{25n}$ from Equation (5) and the baseline neighborhood IGE and AUM estimates from Equation (1) is that the former are constructed after explicitly controlling for selection via $\boldsymbol{W}_{in}$ and K_{in} (under Assumption U). This provides insight of the relative contributions of selection and location effects in heterogeneity in mobility, which will be formally analyzed through a decomposition exercise in the next subsection.

Overview of Results. Figure 3 reports estimates of the three types of AUM effects from Equation (5) for both municipalities and parishes; Appendix Figure A.5 reports estimates of the IGE. We discuss each panel in turn.

The left panels plot the mean neighborhood-level AUM estimates from Equation (5) (under each assumption of selection) against the baseline AUM estimates from Equation (1) through a binscatter plot. This provides a test of our generalized mobility model's fit against the workhorse empirical model.[27] We find that neighborhood-level mobility estimates after controlling for selection are remarkably similar to the baseline model: correlations between the AUM estimates range between 0.93–0.98. Moreover, as seen by the overlap

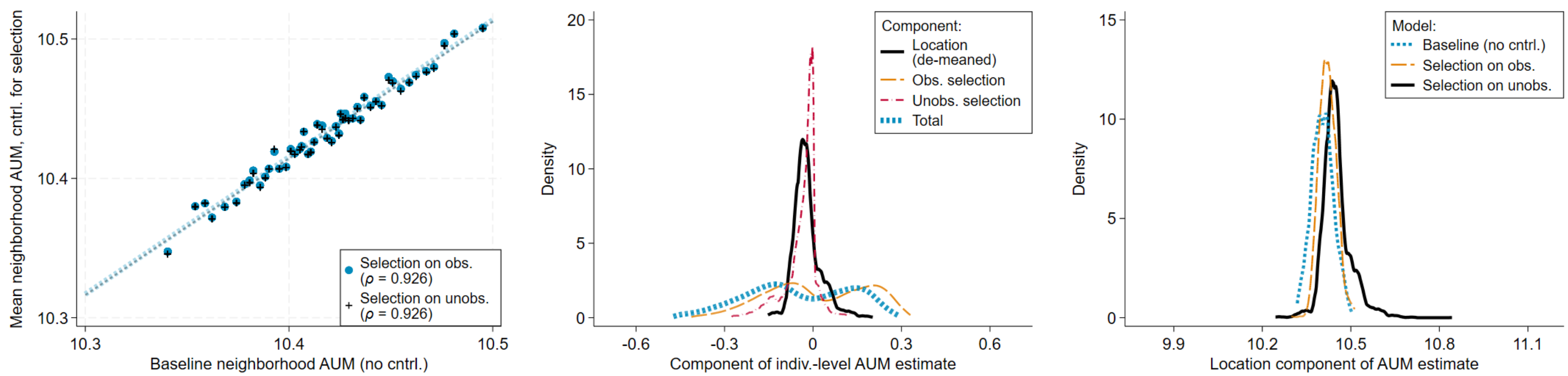
26. Since social characteristics possess relatively smaller effects on child income in practice, we group family and social characteristics together in our analysis.

27. Note that this does not test either model's goodness of fit to the data. We will revisit the limits of our linear specification in Section 5.2.

Figure 3: AUM estimates from the generalized mobility model

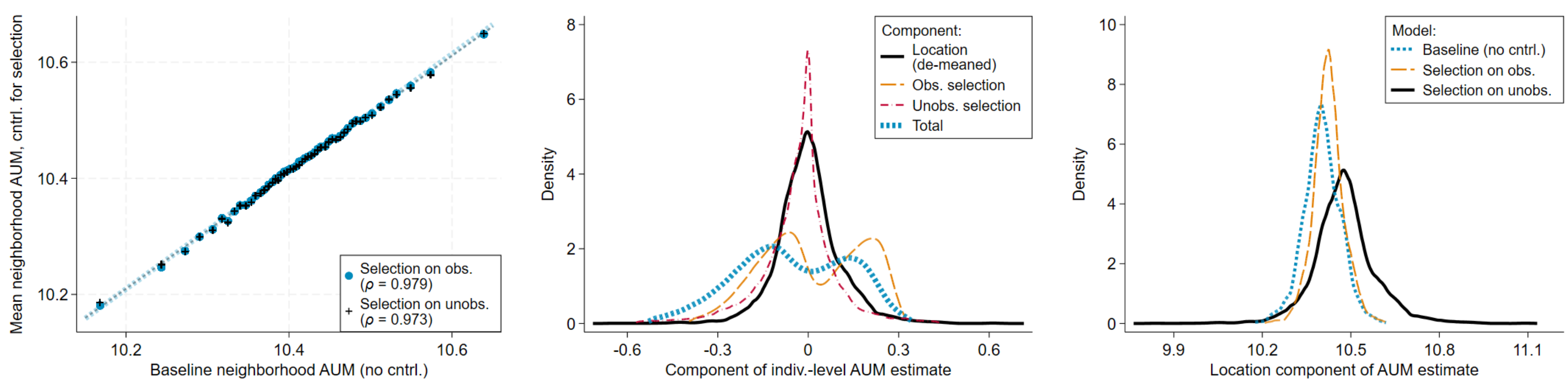
(a) Municipalities

(i) Comparison with baseline estimates (ii) PDFs of indiv.-level AUM components (iii) PDFs of location effects, by model



(b) Parishes

(i) Comparison with baseline estimates (ii) PDFs of indiv.-level AUM components (iii) PDFs of location effects, by model



Notes: Left panels are binscatters of estimates of $\bar{y}^c_{25n}$ from Equations (1) and (5) under Assumptions O and U, binned by two centiles. Dashed lines are lines of best fit. Correlations (ρ) are reported in the legend. Middle panels plot densities of estimates of the individual-level AUM $\bar{y}^c_{i,25n} \equiv \alpha(\boldsymbol{W}_{in}, n) + \beta(\boldsymbol{W}_{in}, n)y^p_{25} + K_{in}$ from Equation (5) and its components under Assumption U. Estimates of the location AUM component $\bar{y}^c_{25n} \equiv \alpha_n + \beta_n y^p_{25}$ are not weighted by neighborhood sample size and de-meaned to facilitate comparisons with the AUM's observable selection component $\boldsymbol{\alpha_w W}_i + \boldsymbol{\beta_w W}_i y^p_{25}$ (since $\boldsymbol{W}_i$ was originally de-meaned). Right panels compare densities of the AUM's location component estimates from Equations (1) and (5) under Assumptions O and U. All densities are weighted by their inverse squared standard errors and Epanechnikov kernels with their respective rule-of-thumb bandwidths, trimmed at the 1st and 99th percentiles to retain a reasonable scale.

of the binscatters, the mean neighborhood-level estimates under Assumptions O and U are nearly identical (correlation > 0.99). These results indicate that both the baseline model and models controlling for selection do an equally good job at predicting average mobility in each neighborhood.

Next, the panels in the middle column present densities of individual-level AUM estimates from the generalized mobility model (under Assumption U) and unpack their underlying components. Despite the strong concordance between the mean neighborhood-level mobility estimates, there is a large degree of heterogeneity in predicted child income across the population that cannot be explained by the pure location component. The empirical density of the observable selection effect component has much larger spread than location effects, and shares similar variation as the overall individual-level AUM distribution. Interestingly, the density of observable effects is bimodal, with one peak at around -0.2 and another peak at around 0.2, suggesting that there are two "types" of family and social characteristic vectors that generate 20% increase or decrease in child income relative to the population average.[28]

Similarly, the empirical density of the selection effect's control function component ranges over negative and positive values. This indicates that there are both positive and negative effects from selection on unobservables, indicating substantial heterogeneity in the effects of unobserved features of families or locations. The fact that many children experience negative effects suggests that their families may have adverse unobserved characteristics associated with their residential preferences and/or face frictions in moving to higher-opportunity neighborhoods (Bergman et al. 2024). Meanwhile, location effect estimates possess substantially smaller variation than the observable selection component, suggesting that the effects of selection could matter more than locations per se. Note, however, that the density of parish location effect estimates has greater spread and more extreme tails than municipalities.

Finally, the right panels compare the density of the location AUM effects across different model specifications. Relative to the baseline AUM estimates from Equation (1), controlling for selection shifts the distribution of location effects to the right. Under Assumption O, controlling for selection decreases its variance across both parishes and municipalities, suggesting that conditioning on observables decreases the variation in the location component estimates. The same is true under Assumption U for municipalities, though the variance of location AUM component estimates increases for parishes.

Though these descriptive patterns are interesting, it is difficult to draw firm assessments about the impact of controlling for selection given the presence of sampling error. We thus pursue a formal variance decomposition that accounts for sampling error in Section 4.3.

28. Recall that $\boldsymbol{W}_{in}$ is de-meaned, which implies the observable selection component has mean zero by construction.

Effects of Family and Social Characteristics. We briefly examine the effects of individual family and social characteristics to better understand the underlying drivers of the observable selection component.

Our model employs an unusually rich vector of family and social characteristics. To simplify our discussion, we assign elements of the covariates $\boldsymbol{W}_{in}$ into eight mutually exclusive categories: parental age, parental assets, mother's employment, parental education, household structure, marital status, parental health, and parental crime. We reverse the sign of any variable with negative effects on child income to set a common direction interpretation. Within each category, we estimate the first principal component (PC) of all individual family characteristics and of all social characteristics. We then construct linear combinations of these family and social characteristics' respective estimates in $(\boldsymbol{\alpha_x}, \boldsymbol{\beta_x})$ and $(\boldsymbol{\alpha_s}, \boldsymbol{\beta_s})$, using their PC loadings as weights, and normalize these linear combinations by their PC scores' standard deviations. This procedure delivers standardized effect estimates of each category of family and social characteristics.

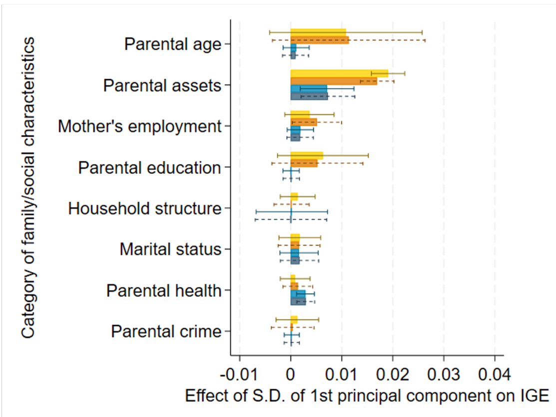
Figure 4 presents our results. In general, individual family characteristics play a far larger role than social characteristics, and effect magnitudes are qualitatively similar between muncipalities and parishes. Controlling for selection on unobservables only modestly attenuates effects on family characteristics. Assets and education of individual families increase the IGE by 0.005–0.02, suggesting a complementarity between parent's own human capital and wealth and their investments in child human capital as predicted by Becker and Tomes (1979). Among the most impactful social characteristics on the IGE is parental assets, which is consistent with Durlauf's (1996) prediction that public goods arising from neighborhood wealth can amplify individual family investments in children. Parental age, which is often controlled for in empirical studies on national-level mobility (e.g., Solon 1992) but rarely in the neighborhood mobility literature, has among the largest effects on upward mobility. This characteristic, along with parental assets, mother's employment, and parental education, increases AUM by 2–3.5%. Household structure, marital status, health, and crime do not affect the IGE but have modest, statistically significant effects on AUM.

The Geography of Mobility, Redux. Next, we take a closer look at the spatial patterns of the location effect estimates that control for selection. Figure 5, Panels (a) and (b) illustrate municipality heat maps of the estimated location effect components of the IGE and AUM across different models. The left maps plot estimates from the baseline model (Equation 1) to provide a reference of spatial patterns while the middle maps plot estimates of the residual location effect components from our generalized mobility model (i.e., based on α_n and β_n from Equation 5) under Assumption U.

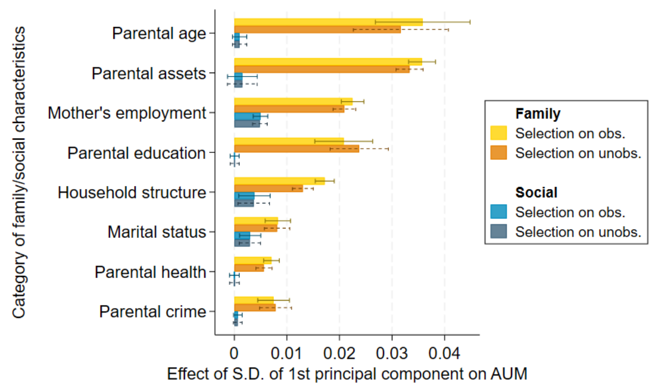
Figure 4: Effects of family and social characteristics on individual-level mobility estimates

(a) Municipalities

(i) Effect on IGE estimates

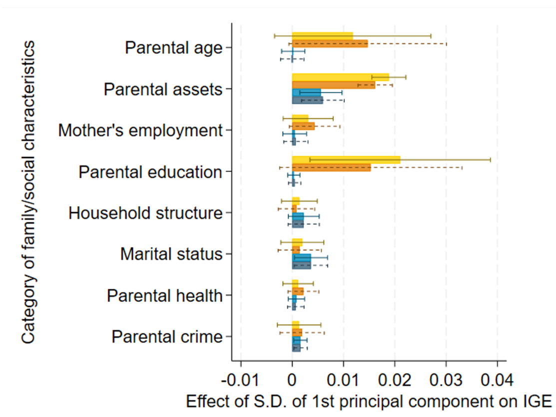


(ii) Effect on AUM estimates

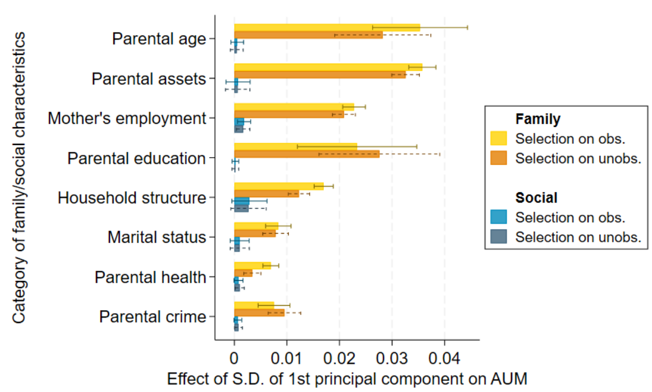


(b) Parishes

(i) Effect on IGE estimates



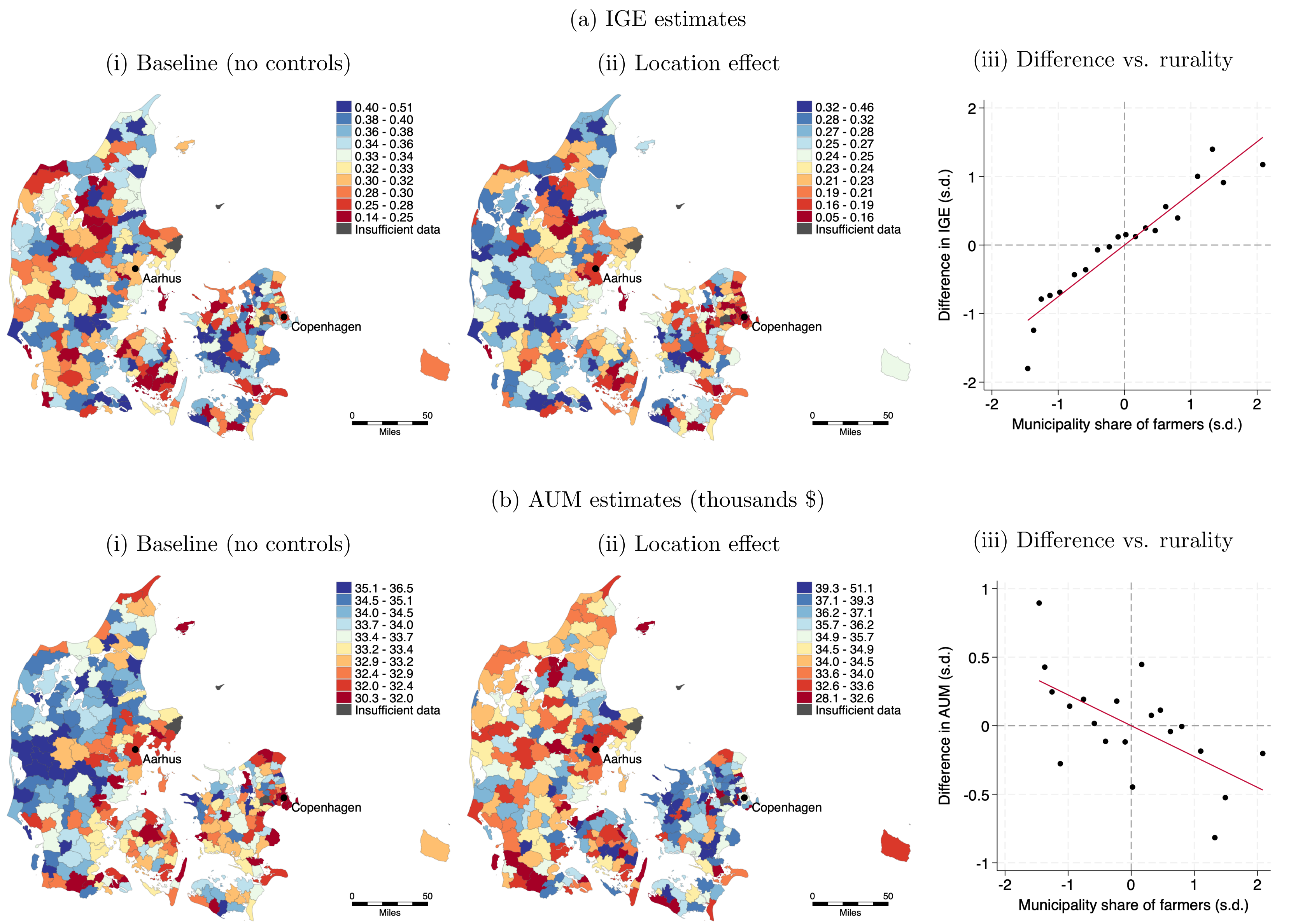
(ii) Effect on AUM estimates



Notes: This figure plots effect sizes of the first principal component of categories of family and social characteristics on the individual-level IGE and AUM estimates, separately based on estimating Equation (5) under Assumption O or U. Principal component loadings of effects are standardized based on the variance of the first principal component score. Categories contain the following variables: "parental age" includes mother and father age; "parental assets" includes sum of mother and father assets; "mother's employment" includes mother's labor force participation along the extensive and intensive margins; "parental education" includes mother and father's years of schooling and indicators of college education; "household structure" includes an indicator and years of childhood in intact families, household size, number of siblings, and number of parental figures; "marital status" includes an indicator and years of childhood in a divorced and in an unmarried household; "parental health" includes an indicator, years of childhood, and duration of parents' inpatient hospitalization; and "parental crime" includes an indicator and years of childhood of parental crime and incarceration. Horizontal capped lines are 90% confidence intervals; solid and dashed lines correspond to Assumptions O and U, respectively.

Remarkably, broad regional patterns in mobility predicted by the workhorse model are reversed after controlling for selection. Municipalities that tend to have greater baseline mo-

Figure 5: Heat maps of location effect estimates, before and after controlling for selection



Notes: This figure plots heat maps of estimates of the location IGE β_n and exponentiated location AUM $\bar{y}^c_{25n}$. Left panels plot estimates from Equation (1). Middle panels plot location effect estimates from Equation (5) under Assumption U. Right panels plot binscatters of standardized differences between the middle and left panels and standardized long-run proportion of farmers in the municipality; data is binned in ventiles and red lines are best linear fits.

bility (higher baseline AUM, lower baseline IGE) lie in much of western and northern Jutland (Denmark's large western peninsula)—which comprise the country's rural mainland—while municipalities with lower baseline mobility are concentrated around large cities such as the greater Copenhagen and Aarhus areas. The opposite pattern, however, is found in heat maps of the residual location effect components. To see this more clearly, the right panels provide binscatter plots between the difference in the municipality-level mobility estimates and the municipality's long-run share of farmers, which proxies rurality. There is a striking association: More rural municipalities tend to possess location effects that diminish mobility.

This result implies that the key engine of intergenerational mobility in rural areas is their self-selected population. Meanwhile, urban areas tend to be endowed with place-based features that promote mobility. More generally, these findings suggest that controlling for selection can dramatically affect our understanding of the spatial landscape of mobility. Employing baseline mobility estimates alone cannot diagnose the relative role of selection or locations in explaining these regional patterns.

4.3 Variance Decomposition

What are the relative contributions of selection and location effects in explaining spatial variation in mobility? To answer this question, we decompose the total variance of the individual-level mobility parameters derived from Equation (5) along these dimensions. First, consider the individual-level IGE:

$$\text{Var}\left(\beta(\widehat{\boldsymbol{W}_{in}, n})\right) = \underbrace{\text{Var}\left(\hat{\boldsymbol{\beta}}_{\boldsymbol{w}}\boldsymbol{W}_{in}\right)}_{\text{Observable selection effect}} + \underbrace{\text{Var}\left(\hat{\beta}_n\right)}_{\text{Location effect}} + \underbrace{2\text{Cov}\left(\hat{\boldsymbol{\beta}}_{\boldsymbol{w}}\boldsymbol{W}_{in}, \hat{\beta}_n\right)}_{\text{Sorting effect}}. \tag{7}$$

The first two terms are variances of the observable selection and location effects defined in Section 4.2. The third term is the "sorting effect," which captures the extent that observable selection effects on mobility are associated with their chosen location's effect. We can further decompose each term *between* and *within neighborhoods*. The between-neighborhood component corresponds to variation in the mean neighborhood-level mobility estimates. Here, we are particularly interested in quantifying the relative contributions of the population and location effects toward between-neighborhood variation. As before, sampling error may explain much of the observed between-neighborhood variation, so we further decompose all terms into their signal and noise components.[29] See Appendix F for the formula of the full decomposition.

Under Assumption O, we can conduct a similar decomposition for $\alpha(\widehat{\boldsymbol{W}_{in}, n})$ and the

29. We assign noise to the "within-neighborhood" component.

covariance Cov $\left(\alpha(\widehat{\boldsymbol{W}_{in}, n}),\ \beta(\widehat{\boldsymbol{W}_{in}, n})\right)$ to decompose the individual-level AUM. Under Assumption U, we consider excluding or including the control function in our decomposition.[30]

Table 2 reports variance decompositions of the individual-level AUM estimates under Assumption O and U for municipalities (Panel A) and parishes (Panel B).[31] All components are expressed as shares of the total estimates to make explicit the relative contributions of each variance component.

There are three key takeaways. First, observable selection effects account for a significant fraction of variation in individual-level AUM estimates. Looking at the top rows of each panel, Column (2), observable selection effects accounts for at least 89% of the variation in the municipality regression model and between 43–91% of the variation in the parish regression model, depending on the model specification.[32] This implies that much of the variation in individual-level mobility is explained by family and social characteristics within neighborhoods rather than location or sorting effects, per se.[33]

Second, the between-neighborhood component (which represents a signal component) explains little of the variation in mobility. Row 1, Column (1) reports that this component makes up at most 9% of the variation, irrespective of neighborhood unit or our modeling assumption on selection.[34]

Third, within the between-neighborhood component, the magnitude of the observable selection effect consistently exceeds that of the location effect. For example, comparing Columns (2) and (3) of Row 1, Assumption O, the between-neighborhood observable selection effect is 1.8–3.2 (5.1% ÷ 2.9% for parishes; 5.5% ÷ 1.7% for municipalities) times larger than the between-neighborhood location effect. This finding is robust to modeling assumptions of selection. This signifies that neighborhood segregation of observable population characteristics generates more variation in mobility than residual location effects. In sum, this analysis demonstrates that most of the variation of mobility experienced by individual families and found across neighborhoods is explained by selection.

30. There are different rationales for excluding or including the control function K_{in}. Excluding K_{in} may be appropriate if K_{in} is viewed as a nuisance parameter that only serves to obtain unbiased estimates of $(\alpha(\boldsymbol{W}_{in}, n),\ \beta(\boldsymbol{W}_{in}, n))$. Alternatively, one may include the control function estimates under the "sorting effect" with the view that selection on unobservables should be accounted for in the mobility process.

31. Appendix Table A.2 reports results for individual-level IGE estimates.

32. Rows 2.1 reports the vast majority of this is explained by pure signal variation *within* neighborhoods.

33. Sorting effect patterns differ across neighborhood units and modeling assumptions. Families endowed with characteristics with larger effects on child income tend to live in municipalities with *lower* location effects, but sort into parishes with *higher* location effects, reflecting the granularity of neighborhood sorting behavior. Including control function estimates generates positive within-neighborhood sorting effects, indicating a positive relationship between the unobserved, idiosyncratic population and location effects.

34. In fact, including the control function drives this component to fall below zero due to strong negative sorting effects.

Table 2: Variance decomposition of individual-level AUM estimates

	Assump.	C.F.	(1) Total (%)	(2) Obs. Selection (%)	(3) Location (%)	(4) Sorting (%)
A. Municipalities						
Total	O		100.0	100.9	2.6	−3.5
	U		100.0	88.8	12.0	−0.7
	U	✓	100.0	96.7	13.0	−9.7
1. Between-neighborhood	O		4.4	5.5	1.7	−3.2
	U		8.1	4.6	3.9	−0.4
	U	✓	−4.0	5.0	4.3	−13.2
2. Within-neighborhood	O		95.6	95.5	0.9	−0.3
	U		91.9	84.2	8.0	−0.3
	U	✓	104.0	91.7	8.8	3.5
2.1. Signal	O		95.0	95.0	—	—
	U		83.8	83.8	—	—
	U	✓	98.1	91.2	—	6.9
2.2. Noise	O		0.6	0.4	0.9	−0.3
	U		8.2	0.4	8.0	−0.3
	U	✓	5.9	0.4	8.8	−3.3
Total variance	O			0.034		
	U			0.037		
	U	✓		0.034		
B. Parishes						
Total	O		100.0	91.2	8.3	0.4
	U		100.0	43.0	56.4	0.6
	U	✓	100.0	77.8	102.2	−80.0
1. Between-neighborhood	O		8.7	5.1	2.9	0.6
	U		2.4	2.4	−0.6	0.7
	U	✓	−91.1	4.4	−1.2	−94.4
2. Within-neighborhood	O		91.3	86.1	5.4	−0.1
	U		97.6	40.5	57.1	−0.1
	U	✓	191.1	73.4	103.4	14.3
2.1. Signal	O		85.8	85.8	—	—
	U		40.4	40.4	—	—
	U	✓	162.6	73.1	—	89.5
2.2. Noise	O		5.6	0.3	5.4	−0.1
	U		57.2	0.2	57.1	−0.1
	U	✓	28.5	0.3	103.4	−75.1
Total variance	O			0.036		
	U			0.074		
	U	✓		0.041		

Notes: This table reports fractions (expressed as percentages) of the total variance of individual-level AUM estimates (from Equation (5)) via the decomposition described in the main text (see Appendix F for details). Column "Assump." indicates whether the decomposition is based on regression estimates under Assumptions O and U. Column "C.F." indicates whether, under Assumption U, control function estimates are included in the sorting effect estimates or excluded from the decomposition.

Accounting for Sampling Error. Having controlled for selection, we next investigate the degree by which residual location effects vary across neighborhoods. In Appendix Tables E.3 and E.5, we conduct analogous hypothesis tests and variance decomposition exercises to those conducted using baseline estimates on the residual location effect estimates from

Equation (5). Overall, these results show that controlling for selection injects significantly more imprecision into the location parameter estimates. Based on hypothesis tests, the majority of neighborhoods' mobility estimates are indistinguishable from estimates implied by a population-level mobility model that controls for selection. In fact, Appendix Table E.3 finds that controlling for selection on unobservables causes all or nearly all of the neighborhoods' location effects to be statistically indistinguishable from their population-level counterparts.

Turning to signal-noise decompositions of the residual location estimates, Appendix Table E.5 reports that between 32–65% of variance of the IGE and AUM location components across municipalities is driven by signal; this fraction is significantly lower, but typically positive, among parishes.[35] Thus, there is still *some positive* signal variance detected in our location effect estimates. Though our hypothesis tests suggest that we cannot draw conclusions about individual neighborhoods, these variance decomposition exercises imply that there is still promise of garnering information about properties of the distribution of estimates, particularly among municipalities. We thus turn to examining the correlates of location effects across neighborhoods.

4.4 Second-Stage Correlations

Figure 6 re-examines the bivariate correlation exercises with location characteristics $\boldsymbol{Z}_n$ using the location AUM effect estimates purged of selection.[36] The square and triangular markers visually depict correlation coefficients under Assumptions O and U respectively. For reference, hollow markers report correlations using the baseline AUM estimates from Equation (1) that do not control for selection.[37] All correlations weigh neighborhoods by their AUM estimates' inverse squared standard error to account for precision.

This exercise provides three key takeaways. First, as predicted in Section 3.2, stripping neighborhood mobility estimates of selection dramatically affects correlation patterns for several location characteristics. Some of the strongest positive predictors of baseline AUM estimates—including the share of homeowners and mean home value—possess little to no association with pure location AUM effect estimates. This suggests that the extent to which baseline neighborhood mobility rates are capitalized into home prices is primarily driven by the composition of families in the neighborhood. Similarly, controlling for selection causes the correlation with mean neighborhood income to vanish, suggesting relatively little role for

35. The only exception is the weighted signal-noise decomposition of parish AUM estimates: As indicated in Row 2.2 of Table E.5, Panel B, the fraction of weighted variance explained by noise actually exceeds 100%. However, the unweighted signal-noise decomposition lowers this share to 93.7%.

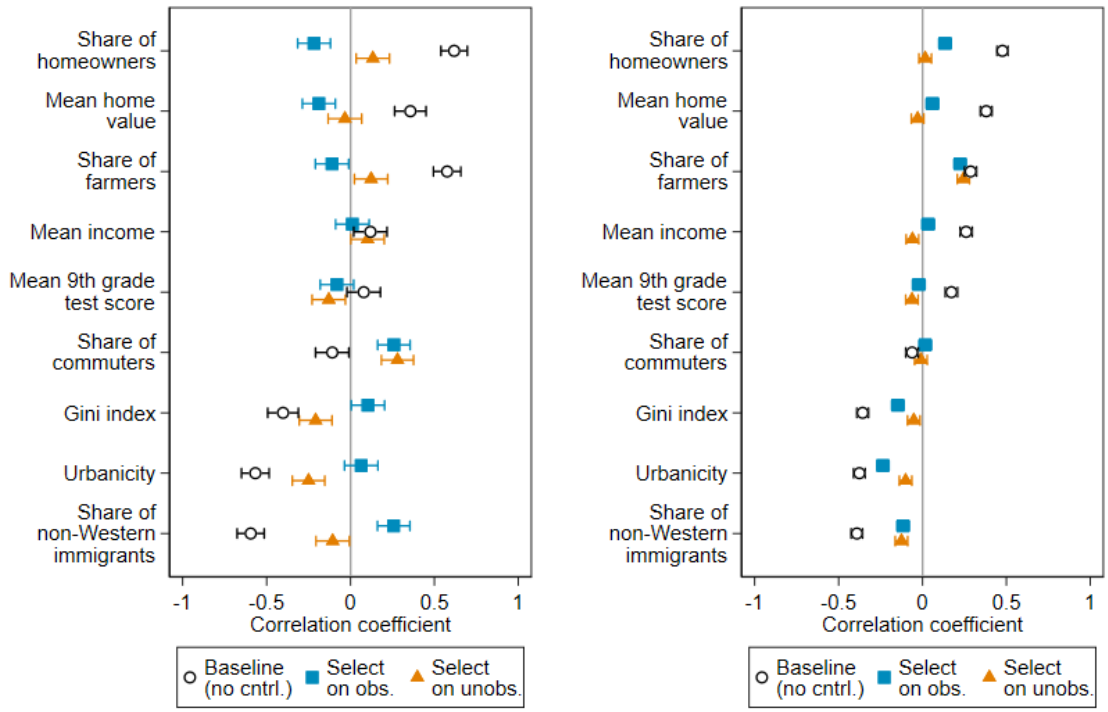
36. Appendix Figure A.6 plots correlations for the location IGE effect estimates.

37. Chetty and Hendren (2018b) conduct a similar second-stage correlation exercise using the location parameters identified via the movers design.

Figure 6: Second-stage correlations between location characteristics and AUM location effect estimates

(a) Municipalities

(b) Parishes



Notes: This figure plots estimates of the correlation coefficients between various location characteristics of neighborhoods and the residual location effect estimates from Equation (5). Square markers are estimates under Assumption O; triangular markers are estimates under Assumption U. Horizontal capped lines are 90% confidence intervals. Hollow markers are correlations with the baseline neighborhood-level AUM estimates from Equation (1), which are included for reference.

place-based amenities such as school quality that may be proxied by income in explaining location effects. This idea is corroborated by the null associations between location effects and mean 9th grade test scores, a remarkable finding that aligns with Wodtke et al. (2023). In other words, the strong positive association between parent income and mean test scores documented in Figure 1(b) implies that schools in Denmark appear to matter for mobility due to composition of families in schools rather than the quality of schooling institutions, per se.

Second, there are some notable differences in correlation patterns found in municipalities compared to parishes. Interestingly, controlling for selection on observables only, compared to additionally controlling for unobservables, affects the magnitude and even the sign of the correlations among municipalities. In contrast, correlations with location AUM estimates among parishes are insensitive to modeling assumptions. This suggests that controlling for selection on unosbervables may be particularly important among coarser geographic units that encapsulate populations that are diverse along unobserved characteristics. Accounting for such selection causes the correlations among municipalities to be more qualitatively similar to those among parishes, though the magnitude of the correlations varies between the neighborhood units.

Third, after controlling for both selection on observables and unobservables, the strongest predictors of location AUM effects signify the neighborhood's rurality (the share of farmers for parishes; urbanicity for municipalities). This is also seen by the strong spatial autocorrelation in location AUM estimates illustrated in Figure 5. It suggests that a key locational source of heterogeneity in neighborhood mobility in Denmark is the difference in amenities in rural versus urban areas. This finding can help refine the appropriate portfolio of place-based policies aimed to promote upward mobility. Our analysis suggests that policies directed to promote place-based investments in amenities found in rural areas may be a promising avenue for improving mobility.

5 Irreducible Heterogeneity across Neighborhoods

The previous bivariate correlation exercise can be naturally extended to a hierarchical linear modeling framework by adjoining Equation (5) with the following "second-level" model (Raudenbush and Bryk 2002):

$$\begin{aligned} \hat{\alpha}_n &= \bar{\alpha} + \bar{\boldsymbol{\alpha}}_{\boldsymbol{z}} \boldsymbol{Z}_n + \alpha_n^{ir} \\ \hat{\beta}_n &= \bar{\beta} + \bar{\boldsymbol{\beta}}_{\boldsymbol{z}} \boldsymbol{Z}_n + \beta_n^{ir}. \end{aligned} \tag{8}$$

Here, the location effect estimates $(\hat{\alpha}_n, \hat{\beta}_n)$ are regressed on the entire vector of location characteristics. The residuals of the model, $(\alpha_n^{ir}, \beta_n^{ir})$, are the *irreducible location effects*. These represent the effect of unknown features of locations that cannot be explained after controlling for both selection and observed location characteristics via Equations (5) and (8). In this sense, irreducible heterogeneity is a good measure of our remaining ignorance of why neighborhoods matter, since the indices themselves are simply labels.

Is there an underlying, interpretable structure in irreducible heterogeneity? To answer this question, this section pursues two extensions. First, we examine if there are clusters underlying the joint distribution of irreducible effects $(\alpha_n^{ir}, \beta_n^{ir})$. Second, we assess whether irreducible heterogeneity might be explained by nonlinear family or social effects present in the intergenerational transmission of income.

As indicated in Appendix Tables E.3 and E.5, the underlying location effect estimates $(\hat{\alpha}_n, \hat{\beta}_n)$ are extremely noisy measures of the true location effects (α_n, β_n). This motivates applying empirical Bayes (EB) shrinkage methods to develop biased but more precise (in the mean squared error sense) estimators of irreducible location effects to utilize in our two extensions. We summarize our EB procedure below.

Empirical Bayes Estimates. In the case of $\hat{\beta}_n$ (and analogously for $\hat{\alpha}_n$), we assume that $\hat{\beta}_n \sim N(\beta_n,\ \sigma^2_{\hat{\beta}_n})$ and specify that the true residual location effect parameter β_n itself is drawn from the prior hyperdistribution $\beta_n \sim N(\bar{\beta} + \bar{\boldsymbol{\beta}}_{\boldsymbol{z}} \boldsymbol{Z}_n,\ \nu^2_\beta)$. This implies the following EB posterior mean:

$$\hat{\beta}_n^{EB} \equiv \tfrac{\nu^2_\beta}{\nu^2_\beta + \sigma^2_{\hat{\beta}_n}} \cdot \hat{\beta}_n + \tfrac{\sigma^2_{\hat{\beta}_n}}{\nu^2_\beta + \sigma^2_{\hat{\beta}_n}} \cdot (\bar{\beta} + \bar{\boldsymbol{\beta}}_{\boldsymbol{z}} \boldsymbol{Z}_n) \tag{9}$$

The variance $\sigma^2_{\hat{\beta}_n}$ is estimated by the squared standard errors of $\hat{\beta}_n$. We follow the procedure described in Morris (1983) to estimate the hyperparameters. Intuitively, the shrinkage factor (i.e., $\frac{\hat{\nu}^2_\beta}{\hat{\nu}^2_\beta + \hat{\sigma}^2_{\hat{\beta}_n}}$) adjusts such that the noisier the original least-squares estimator (i.e., larger $\hat{\sigma}^2_{\hat{\beta}_n}$), the more the EB posterior mean will shrink the original estimator $\hat{\beta}_n$ toward the predicted value of the second-level model (i.e., $\hat{\bar{\beta}} + \hat{\bar{\boldsymbol{\beta}}}_{\boldsymbol{z}} \boldsymbol{Z}_n$).

Substituting Equation (8) into the posterior means, the EB estimators of the irreducible location effect components are simply

$$\begin{aligned} \hat{\alpha}_n^{ir,EB} &= \tfrac{\hat{\nu}^2_\alpha}{\hat{\nu}^2_\alpha + \hat{\sigma}^2_{\hat{\alpha}_n}} \cdot \hat{\alpha}_n^{ir} \\ \hat{\beta}_n^{ir,EB} &= \tfrac{\hat{\nu}^2_\beta}{\hat{\nu}^2_\beta + \hat{\sigma}^2_{\hat{\beta}_n}} \cdot \hat{\beta}_n^{ir}. \end{aligned} \tag{10}$$

We will work with these EB posterior mean estimates to analyze the structure of irreducible heterogeneity separately in municipalities and parishes. In what follows, we focus exclusively on estimates derived under Assumption U.

5.1 Neighborhood Types

Our first extension aims to identify latent "types" of neighborhoods in the joint distribution of irreducible location effect estimates. To investigate this, we conduct a k-means analysis on $(\hat{\alpha}_n^{ir,EB}, \hat{\beta}_n^{ir,EB})$. The k-means algorithm assigns neighborhoods to K clusters such that the assignments minimize the within-cluster sum of squares (WSS) of distances between $(\hat{\alpha}_n^{ir,EB}, \hat{\beta}_n^{ir,EB})$ and their respective cluster means. Using standard rule-of-thumb assessments based on the WSS, we find that neighborhoods fall into $K = 2$ latent types.[38]

Table 3, Panel A reports differences in the irreducible effects between the neighborhood

38. See Appendix Figure A.7 for scree plots of WSS and the proportional reduction of error for different choices of K. Appendix Table A.3 provides results for choosing $K = 3$.

types. For both parishes and municipalities, Type 1 is characterized by lower $\hat{\alpha}_n^{ir,EB}$ and higher $\hat{\beta}_n^{ir,EB}$ estimates than Type 2. Specifically, on average, Type 1's $\hat{\beta}_n^{ir,EB}$ is between 0.02–0.05 higher than Type 2. Meanwhile, its relatively smaller $\hat{\alpha}_n^{ir,EB}$ offsets this difference, resulting in a 0.4–0.5% lower EB irredicuble location AUM effect than Type 2. Qualitatively similar patterns hold for the unshrunken irreducible location effects $(\hat{\alpha}_n^{ir}, \hat{\beta}_n^{ir})$ and the location effects $(\hat{\alpha}_n, \hat{\beta}_n)$. There are less consistent patterns in the population or control function components of the IGE or AUM between municipality and parish types.

Panel B turns to differences in mean population characteristics between the neighborhood types. Interestingly, we find that there are statistically significant differences between neighborhood types along a number of dimensions. The presence of significant differences may be surprising given that the neighborhood types are constructed from the irreducible location effect estimates, which have already controlled for these characteristics at both the family and social levels. This reveals the potential limits of the linear model specifications employed thus far.

5.2 Nonlinearities in Population Characteristics

All of our main results have been derived from Equation (5), which is a linear specification of Equation (2) described in the introduction. In Equation (2), $\alpha(\cdot)$ and $\beta(\cdot)$ are general functions that may nonlinearly depend on family characteristics $\boldsymbol{X}_i$ and social characteristics $\boldsymbol{S}_{-in}$, as posited by Becker et al. (2018) and Durlauf and Seshadri (2018). Our second extension shows how irreducible location effects may capture nonlinearities in the mobility process.

Let $D_{ij} \equiv \mathbb{1}\{n(i) = j\}$ indicate if family i resides in neighborhood j and let $\tilde{Y}_{ij}^p \equiv D_{ij} \cdot Y_i^p$. Further, denote $f_{ij} \equiv \alpha(\boldsymbol{X}_i, \boldsymbol{S}_{-ij}, j) + \beta(\boldsymbol{X}_i, \boldsymbol{S}_{-ij}, j) \cdot Y_i^p$. Finally, for a given neighborhood n, let Y_i^{c*}, $\tilde{Y}_{in}^{p*}$, f_{in}^* be population residuals from linearly projecting Y_i^c, $\tilde{Y}_{in}^p$, f_{in} onto all other covariates $(\boldsymbol{X}_i,\ \boldsymbol{X}_i \cdot Y_i^p,\ \boldsymbol{S}_{-in},\ \boldsymbol{S}_{-in} \cdot Y_i^p,\ (D_{ij})_{j\in\mathcal{N}},\ (\tilde{Y}_{ij}^p)_{j\neq n})$. If the nonlinear model described in Equation (2) is the true data generating process, then we can invoke the Frisch-Waugh-Lovell theorem to express the location IGE effect as

$$\beta_n = \frac{\mathbb{E}[\tilde{Y}_{in}^{p*} \cdot Y_i^{c*}]}{\mathbb{E}[\tilde{Y}_{in}^{p*2}]} = \frac{\mathbb{E}[\tilde{Y}_{in}^{p*} \cdot f_{in}^*]}{\mathbb{E}[\tilde{Y}_{in}^{p*2}]} \tag{11}$$

The final equality in (11) shows that the identified location-specific component of the IGE is simply an average slope of the residual nonlinear location-specific function f_{in}^* (White 1980). In our context, this captures a combination of (i) pure location effects captured by the neighborhood index n and (ii) neighborhood-level averages of nonlinear effects arising

Table 3: Differences between neighborhood types

	(1)	(2)	(3)	(4)
	Municipalities		Parishes	
	Type 1 Mean	Type 2 Diff.	Type 1 Mean	Type 2 Diff.
A. Mobility estimates				
Irreducible location IGE	0.033	−0.073***	0.081	−0.158***
		(0.004)		(0.004)
E.B. irreducible location IGE	0.009	−0.024***	0.023	−0.049***
		(0.001)		(0.001)
Irreducible location AUM	−0.008	0.023***	−0.032	0.056***
		(0.007)		(0.009)
E.B. irreducible location AUM	−0.003	0.005***	−0.002	0.004***
		(0.002)		(0.000)
Location component of IGE	0.262	−0.080***	0.291	−0.157***
		(0.005)		(0.004)
Location component of AUM	10.453	0.021***	10.435	0.055***
		(0.008)		(0.009)
Population component of IGE	−0.006	0.014***	−0.000	0.000
		(0.003)		(0.001)
Population component of AUM	0.008	−0.018***	0.001	−0.002
		(0.005)		(0.002)
Control function estimate	−0.038	−0.005	−0.029	−0.019**
		(0.008)		(0.009)
B. Mean family characteristics				
Mean mother age	26.835	−0.074	26.815	−0.028
		(0.065)		(0.036)
Mean parent income (thousands)	71.520	−0.274	71.303	0.232
		(1.056)		(0.502)
Mean parent assets (thousands)	25.881	−2.121	25.068	−0.274
		(1.617)		(0.815)
Share of mothers not employed	0.408	0.004	0.411	−0.004
		(0.007)		(0.004)
Mean mother years of education	12.836	0.102	12.875	0.013
		(0.065)		(0.033)
Mean father years of education	13.130	0.104	13.173	0.005
		(0.078)		(0.039)
Share of intact families	0.687	−0.041***	0.673	−0.010*
		(0.011)		(0.005)
Mean household size	4.144	−0.103***	4.110	−0.027**
		(0.027)		(0.012)
Share of divorced families	0.204	0.024***	0.212	0.006*
		(0.007)		(0.004)
Share of never married families	0.142	0.018***	0.147	0.006***
		(0.003)		(0.002)
Share of parents hospitalized	0.730	−0.007*	0.728	−0.003
		(0.004)		(0.002)
Share of parents committed crime	0.138	0.015***	0.142	0.006**
		(0.004)		(0.003)
Share of parents incarcerated	0.055	0.004*	0.056	0.001
		(0.002)		(0.001)
Number of neighborhoods	133	140	768	1,181

Notes: This table reports differences in mean mobility estimates and mean family characteristics between the neighborhood types identified in the k-means clustering algorithm on $(\hat{\alpha}_n^{ir,EB}, \hat{\beta}_n^{ir,EB})$. All estimates are weighted by neighborhood sample size.

from complementarities between Y_i^p and $\boldsymbol{X}_i$ or $\boldsymbol{S}_{-in}$ predicted by economic theory.

From this vantage point, the irreducible components $(\alpha_n^{ir}, \beta_n^{ir})$ can provide suggestive evidence of nonlinearities. The second-level model regressions (8) are purged of observable location characteristics $\boldsymbol{Z}_n$ that are determinants of pure location effects. Thus, the irreducible components capture any local nonlinear effects as well as effects of unobserved location characteristics that cannot be fully explained by $\boldsymbol{Z}_n$. This implies that we expect the irreducible components themselves to be associated with average characteristics of the neighborhood's population, $\bar{\boldsymbol{X}}_n$. This analysis can help adjudicate which dimensions of characteristics can generate nonlinear mobility patterns. For example, if the irreducible components are significantly associated with average mother's education level, then this would lend credence to Becker et al. (2018).

To detect the presence of nonlinearities in the mobility process, we estimate the following regression model,

$$\hat{\beta}_n^{ir,EB} = \gamma + \boldsymbol{\eta}\tilde{\boldsymbol{X}}_n + \xi_n, \tag{12}$$

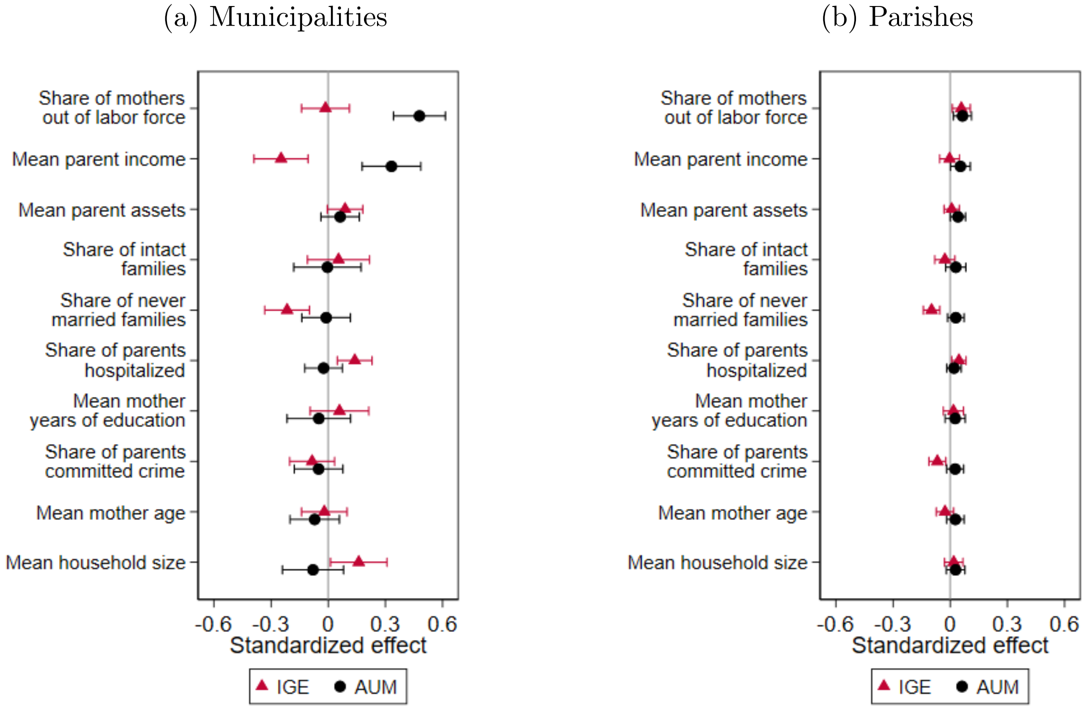
and similarly for $\hat{\bar{y}}_{25n}^{c,ir,EB}$, weighting neighborhoods by their sample sizes. Here, $\tilde{\boldsymbol{X}}_n$ is a vector of neighborhood-level mean family characteristics from our main sample that are residualized by location characteristics $\boldsymbol{Z}_n$ (in accordance with Frisch-Waugh-Lovell logic).[39] We choose a representative subset of the full vector $\boldsymbol{X}_i$ in order to capture distinct dimensions of populations. We re-scale coefficients by the ratio of standard deviations of the outcome variable and regressors so that the $\boldsymbol{\eta}$ coefficients can be compared along a common scale.

Figure 7 reports the resulting estimates. Three interesting patterns emerge. First, there are notable differences in the characteristics that drive irreducible location IGEs (triangular markers) versus AUMs (circular markers). The population characteristics with the most significant associations with the irreducible location IGE effect across parishes are shares of children whose mothers were ever out of the labor force during childhood and whose parents were hospitalized, committed crime, or unmarried. This suggests nonlinearities in the complementarity between these characteristics and parental income in the intergenerational transmission process.

Second, the only characteristics significantly positively associated with irreducible AUM, particularly among municipalities, are mean parental income and the share of mothers out of the labor force. All other characteristics have largely null or small, statistically insignificant associations. Notably, we do not find empirical evidence supporting Becker et al. (2018), since mean mother's years of education has statistically insignificant estimates. In contrast, the large positive effects from parental income and the share of mothers out of the labor

39. Recall that $\boldsymbol{Z}_n$ are defined based on place-based features of the neighborhoods or long-run characteristics of the full population residing in the neighborhood—not from our main sample of families.

Figure 7: Detecting nonlinearities: Results from regressing irreducible location effects on mean family characteristics



Notes: This figure plots estimates from Equation (12), rescaled to standardized effect units. All mean family characteristics are residualized by location characteristics and regressions are weighted by neighborhood sample sizes. Horizontal capped lines are 90% confidence intervals.

force suggest that monetary and maternal time investments may have convex effects in the mobility process (Del Boca et al. 2014). This discussion illustrates that examining the sources of variation of irreducible location heterogeneity through average family characteristics can help to reveal sources of potential nonlinearities, and to refine the set of plausible theories of intergenerational mobility.

Third, there are differences in the relevance of certain population characteristics between municipalities and parishes. For example, a one standard deviation increase in parental income is associated with nearly 0.25 standard deviation decline in the irreducible location IGE effect for municipalities but has a precise null association for parishes. Differences in correlation patterns across neighborhood types suggests that nonlinearities in the mobility process may manifest at certain geographic levels but not others, suggesting the mobility process may operate in complex ways at different levels of neighborhood aggregation.

5.3 Discussion

We conclude this section by noting that the patterns we find in our clustering and nonlinearity analyses are also consistent with more complex models than the ones we study. Among the explanations that can produce these patterns are bimodality in the density of unobserved neighborhood-level variables or the presence of multiple equilibria in the behaviors of children, which can map mean family characteristics to their future income. An example of an unobserved neighborhood-level variable that may be captured in our irreducible location effects is self-efficacy (Sampson et al. 1997; Sampson et al. 1999; Sampson and Raudenbush 1999, 2004), which refers to norms in a neighborhood that involve cooperative behaviors. Models of multiple equilibria such as Brock and Durlauf (2001) produce finite numbers of self-consistent choices for interdependent behaviors, which can produce variation in our neighborhood mobility measures. Such models are generically nonlinear, which can produce different linear approximations for intergenerational mobility relationships at the different equilibria. These differences could, in principle, generate the evidence of distinct neighborhood "types" and nonlinearities we observe. We certainly do not claim to have corroborated either of these more complex neighborhood mobility models. Rather, we emphasize that a systematic understanding of the irreducible heterogeneity we have uncovered may involve very different models than what that is conventional in empirical mobility analysis.

6 Conclusion

In this paper, we develop and estimate a generalized model of intergenerational income mobility to understand the sources of heterogeneity in mobility estimates across neighborhoods. Using administrative data from Denmark, we extend the workhorse neighborhood mobility model by explicitly controlling for selection into 273 municipalities and nearly 2,000 parishes. Our model conditions on rich vectors of family and social characteristics and, in some specifications, we employ a multinomial neighborhood selection model to account for selection on unobservables.

We find that selection plays a predominant role in explaining the heterogeneity in the overall distribution of mobility within population as well as between neighborhoods. We also find that the spatial landscape of mobility and the location characteristics correlated with conventional mobility estimates change dramatically after accounting for selection. This highlights how correlation exercises between conventional estimates and neighborhood characteristics are difficult to interpret and connect to the underlying mechanisms of neighborhood mobility.

Nevertheless, even after accounting for estimation error, location effects still explain some of the neighborhood heterogeneity in mobility. Moreover, after controlling for all observed location characteristics, there remains irreducible neighborhood heterogeneity, which represents the limits of our linear specification of the mobility process. Clustering neighborhoods using irreducible location effects suggests the existence of two underlying "types" of neighborhoods. We also show these irreducible effects can provide a lens for detecting nonlinearities in the mobility process, and find evidence that mean parental income and mothers' labor force participation are candidate factors that nonlinearly affect child income in Denmark.

Our paper complements and extends a range of efforts in the literature that aim to disentangle pure location effects from selection and its underlying mechanisms. Though the specific empirical results documented here may be unique to the Danish context, our framework underscores the importance of considering the role of selection into neighborhoods and provides a template for understanding the sources of heterogeneity in neighborhood mobility around the world.

References

Abdulkadiroğlu, Atila, Parag A. Pathak, Jonathan Schellenberg, and Christopher R. Walters. 2020. "Do Parents Value School Effectiveness?" *American Economic Review* 110 (5): 1502–1539.

Acciari, Paolo, Alberto Polo, and Giovanni L. Violante. 2022. "And Yet It Moves: Intergenerational Mobility in Italy." *American Economic Journal: Applied Economics* 14 (3): 118–163.

Alesina, Alberto, Sebastian Hohmann, Stelios Michalopoulos, and Elias Papaioannou. 2021. "Intergenerational Mobility in Africa." *Econometrica* 89 (1): 1–35.

Andrews, Isaiah, Toru Kitagawa, and Adam McCloskey. 2024. "Inference on Winners." *The Quarterly Journal of Economics* 139 (1): 305–358.

Aydemir, Abdurrahman B., and Hakki Yazici. 2019. "Intergenerational Education Mobility and the Level of Development." *European Economic Review* 116:160–185.

Becker, Gary S., Scott Duke Kominers, Kevin M. Murphy, and Jörg L. Spenkuch. 2018. "A Theory of Intergenerational Mobility." *Journal of Political Economy* 126 (S1): S7–S25.

Becker, Gary S., and Nigel Tomes. 1979. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy* 87 (6): 1153–1189.

Bell, Brian, Jack Blundell, and Stephen Machin. 2022. "Where is the Land of Hope and Glory? The Geography of Intergenerational Mobility in England and Wales." *Scandinavian Journal of Economics.*

Benjamini, Yoav, and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society: Series B (Methodological)* 57 (1): 289–300.

Bergman, Peter, Raj Chetty, Stefanie DeLuca, Nathaniel Hendren, Lawrence F. Katz, and Christopher Palmer. 2024. "Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice." *American Economic Review* 114 (5): 1281–1337.

Blume, Lawrence E., William A. Brock, Steven N. Durlauf, and Yannis M. Ioannides. 2011. "Identification of Social Interactions." In *Handbook of Social Economics,* 1:853–964. Elsevier.

Bourguignon, François, Martin Fournier, and Marc Gurgand. 2007. "Selection Bias Corrections based on the Multinomial Logit Model: Monte Carlo Comparisons." *Journal of Economic Surveys* 21 (1): 174–205.

Britto, Diogo G. C., Alexandre de Andrade Fonseca, Paolo Pinotti, Breno Sampaio, and Lucas Warwar. 2022. "Intergenerational Mobility in the Land of Inequality." CESifo Working Paper 10004.

Brock, William A., and Steven N. Durlauf. 2001. "Discrete Choice with Social Interactions." *The Review of Economic Studies* 68 (2): 235–260.

Chamberlain, Gary. 1986. "Asymptotic Efficiency in Semi-Parametric Models with Censoring." *Journal of Econometrics* 32 (2): 189–218.

Chetty, Raj, John N. Friedman, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter. 2018. "The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility." NBER Working Paper No. 32697, National Bureau of Economic Research.

Chetty, Raj, and Nathaniel Hendren. 2018a. "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects." *The Quarterly Journal of Economics* 133 (3): 1107–1162.

———. 2018b. "The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates." *The Quarterly Journal of Economics* 133 (3): 1163–1228.

Chetty, Raj, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter. 2020. "Race and Economic Opportunity in the United States: An Intergenerational Perspective." *The Quarterly Journal of Economics* 135 (2): 711–783.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *The Quarterly Journal of Economics* 129 (4): 1553–1623.

Cholli, Neil A., and Steven N. Durlauf. 2022. "Intergenerational Mobility." NBER Working Paper 29760, National Bureau of Economic Research.

Chuard, Patrick, and Veronica Grassi. 2020. "Switzer-Land of Opportunity: Intergenerational Income Mobility in the Land of Vocational Education." SSRN Working Paper 3662560.

Connolly, Marie, Miles Corak, and Catherine Haeck. 2019. "Intergenerational Mobility Between and Within Canada and the United States." *Journal of Labor Economics* 37 (S2): S595–S641.

Corak, Miles. 2013. "Income Inequality, Equality of Opportunity, and Intergenerational Mobility." *Journal of Economic Perspectives* 27 (3): 79–102.

———. 2020. "The Canadian Geography of Intergenerational Income Mobility." *The Economic Journal* 130 (631): 2134–2174.

Dahl, Gordon B. 2002. "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets." *Econometrica* 70 (6): 2367–2420.

Damm, Anna Piil, and Marie Louise Schultz-Nielsen. 2008. "Danish Neighbourhoods: Construction and Relevance for Measurement of Residential Segregation." *Nationalokonomisk Tidsskrift* 146 (3): 241–262.

Del Boca, Daniela, Christopher Flinn, and Matthew Wiswall. 2014. "Household Choices and Child Development." *Review of Economic Studies* 81 (1): 137–185.

Deutscher, Nathan. 2020. "Place, Peers, and the Teenage Years: Long-Run Neighborhood Effects in Australia." *American Economic Journal: Applied Economics* 12 (2): 220–249.

Deutscher, Nathan, and Bhashkar Mazumder. 2020. "Intergenerational Mobility across Australia and the Stability of Regional Estimates." *Labour Economics* 66:101861.

Dodin, Majed, Sebastian Findeisen, Lukas Henkel, Dominik Sachs, and Paul Schüle. 2024. "Social Mobility in Germany." *Journal of Public Economics* 232:105074.

Dubin, Jeffrey A., and Daniel L. McFadden. 1984. "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption." *Econometrica: Journal of the Econometric Society,* 345–362.

Durlauf, Steven N. 1996. "A Theory of Persistent Income Inequality." *Journal of Economic Growth* 1:75–93.

Durlauf, Steven N., and Ananth Seshadri. 2018. "Understanding the Great Gatsby Curve." *NBER Macroeconomics Annual* 32 (1): 333–393.

Eriksen, Jesper, and Martin D Munk. 2020. "The Geography of Intergenerational Mobility—Danish Evidence." *Economics Letters* 189:109–24.

Gallagher, Ryan, Robert Kaestner, and Joseph Persky. 2019. "The Geography of Family Differences and Intergenerational Mobility." *Journal of Economic Geography* 19 (3): 589–618.

Güell, Maia, Michele Pellizzari, Giovanni Pica, and José V Rodríguez Mora. 2018. "Correlating Social Mobility and Economic Outcomes." *The Economic Journal* 128 (612): F353–F403.

Heckman, James. 1990. "Varieties of Selection Bias." *The American Economic Review* 80 (2): 313–318.

Heckman, James, and Salvador Navarro. 2004. "Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models." *Review of Economics and Statistics* 86, no. 1 (February): 30–57.

Heckman, James J., and Richard Robb. 1985. "Alternative Methods for Evaluating the Impact of Interventions: An Overview." *Journal of Econometrics* 30 (1): 239–267.

Heidrich, Stefanie. 2017. "Intergenerational Mobility in Sweden: A Regional Perspective." *Journal of Population Economics* 30 (4): 1241–1280.

Kenedi, Gustave, and Louis Sirugue. 2023. "Intergenerational Income Mobility in France: A Comparative and Geographic Analysis." *Journal of Public Economics* 226:104974.

Krueger, Alan B. 2012. "The Rise and Consequences of Inequality in the United States." Speech at the Center for American Progress.

Laliberté, Jean-William. 2021. "Long-term contextual effects in education: Schools and neighborhoods." *American Economic Journal: Economic Policy* 13 (2): 336–377.

Mazumder, Bhashkar. 2005. "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States using Social Security Earnings Data." *Review of Economics and Statistics* 87 (2): 235–255.

Mitnik, Pablo A., and David B. Grusky. 2020. "The Intergenerational Elasticity of What? The Case for Redefining the Workhorse Measure of Economic Mobility." *Sociological Methodology* 50 (1): 47–95.

Mogstad, Magne, Joseph P. Romano, Azeem M. Shaikh, and Daniel Wilhelm. 2024. "Inference for Ranks with Applications to Mobility across Neighbourhoods and Academic Achievement across Countries." *Review of Economic Studies* 91 (1): 476–518.

Morris, Carl N. 1983. "Parametric Empirical Bayes Inference: Theory and Applications." *Journal of the American Statistical Association* 78 (381): 47–55.

Otero, Sebastián, Nano Barahona, and Cauê Dobbin. 2021. "Affirmative Action in Centralized College Admission Systems: Evidence from Brazil." Unpublished manuscript.

Raudenbush, Stephen W., and Anthony S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods.* Vol. 1. SAGE Publications.

Rothbaum, Jonathan. 2016. "Sorting and Geographic Variation in Intergenerational Mobility." Unpublished manuscript.

Rubin, Donald B. 1981. "The Bayesian Bootstrap." *The Annals of Statistics,* 130–134.

Sampson, Robert J., Jeffrey D. Morenoff, and Felton Earls. 1999. "Beyond Social Capital: Spatial Dynamics of Collective Efficacy for Children." *American Sociological Review* 64 (5): 633–660.

Sampson, Robert J., and Stephen W. Raudenbush. 1999. "Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods." *American Journal of Sociology* 105 (3): 603–651.

———. 2004. "Seeing Disorder: Neighborhood Stigma and the Social Construction of 'Broken Windows'." *Social Psychology Quarterly* 67 (4): 319–342.

Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls. 1997. "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *Science* 277 (5328): 918–924.

Solon, Gary. 1992. "Intergenerational Income Mobility in the United States." *The American Economic Review,* 393–408.

White, Halbert. 1980. "Using Least Squares to Approximate Unknown Regression Functions." *International Economic Review,* 149–170.

Wodtke, Geoffrey T., Ugur Yildirim, David J. Harding, and Felix Elwert. 2023. "Are Neighborhood Effects Explained by Differences in School Quality?" *American Journal of Sociology* 128 (5): 1472–1528.