Move on Up?
Ranking Places by Intergenerational Mobility and Revealed Preference*

John Klopfer and Corbin Miller

March 27, 2026

PENDING APPROVAL: DO NOT DISSEMINATE

Abstract

Can places with high intergenerational mobility be ranked and recommended to families, to improve children's outcomes in adulthood? Critics of related policies observe that rankings are based on estimates, and may be completely uninformative because of bias or noise in the underlying estimates. In this paper, we improve quasi-experimental estimates of the exposure effects of commuting zones (CZs) on children's adult outcomes in the U.S., addressing selection bias and reducing estimation noise by two orders of magnitude. Measuring bias in observed outcomes by comparison with our estimates, we show that rankings based on observed outcomes are unreliable. We provide a more reliable ranking of CZs based on exposure effects, which confirms significant differences between CZs under stringent statistical criteria, and could be used to inform families about where and when to move. However, parents prefer places with negative exposure effects. Parents' preferences are explained by earnings premia, setting up a key tradeoff for policies encouraging family relocation. Finally, we explore heterogeneity, finding that native-born children of native or immigrant parents have similar exposure effects.

*Klopfer: Faculty of Business and Economics, The University of Hong Kong; Miller: Office of Research, Applied Analytics, and Statistics, Internal Revenue Service. Feedback is appreciated. Correspondence: johnklopfer@gmail.com

We thank Philipp Barteska, Matthias Fahn, Naijia Guo, Jasmine Hao, Alan Kwan, Gedeon Lim, Juan Pantano, Uta Schoenberg, Ashley Wong, Michael Wong, Liang Zhong, and participants at the MEA 2026 Annual Conference SOLE Sessions for helpful discussions; Klopfer thanks the Hoover Institution for its hospitality, and the University of Hong Kong and the Research Grants Council of Hong Kong (GRF 17500922) for their support. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors and do not reflect the views or the official positions of the IRS or any of our colleagues.

1 Introduction

Children's outcomes vary across places (Chetty et al., 2014 and 2026), motivating interest in policies and programs that relocate families to "better" places, for example HUD's Moving to Opportunity (MTO) and Seattle-King County's Creating Moves to Opportunity (CMTO) demonstrations (see Chetty, Hendren, and Katz, 2016 and Bergman et al., 2024). Making recommendations for where to relocate requires ranking places, and ranking places requires estimates that are free of *both* bias *and* noise. While estimates of observed outcomes are precise, and there is evidence that they contain a causal signal (Chetty and Hendren, 2018a; Chetty et al. 2026), it remains unclear how much of the variation in observed outcomes is bias. Further, while quasi-experimental estimates of exposure effects show great promise in removing bias (Chetty and Hendren, 2018b), most of the variation in exposure effects is noise, making the estimates too imprecise to measure bias and their ranks too uninformative to make recommendations (Mogstad et al., 2024; Andrews et al., 2024).

Which places give children the greatest chance at upward mobility, and can we rank them with any confidence? To answer these questions, we work to improve quasi-experimental estimates of childhood exposure effects on children's adult outcomes, for each commuting zone (CZ) in the United States. We revisit the exposure-time design developed by Chetty and Hendren (2018b), asking whether the dimensionality of the benchmark estimator can be reduced without sacrificing efficiency, and carrying out stronger tests of the identifying assumption. Then, we ask whether quasi-experimental methods are necessary, comparing our estimates with observed outcomes to measure bias in the observed outcomes. With evidence that our estimates are the most efficient unbiased estimates available, we rank CZs under increasingly stringent statistical criteria. Finally, we address practical questions for policies based on these estimates, including the alignment between rankings and parents' preferences, and between rankings for heterogeneous sub-populations.

The fundamental challenge for estimates of exposure effects is the dimensionality of the causal model: because parents can move between any pair of CZs and each move represents a choice that may be related to unobservables, the most credible estimates for CZs must control for potentially millions of move-specific fixed effects. Yet, the low frequency of many underlying moves leaves many of the fixed effects un-identified, and the size of the remaining design matrix makes the regression computationally infeasible. Chetty and Hendren (2018b) tackled this problem by breaking it into the roughly 12,000 most-common moves to estimate pairwise exposure effects, before recombining those moves into weighted-average estimates for the roughly 600 most-populous CZs. Unfortunately, while these choices made estimation feasible, they create the noise critics object to: each exposure effect is the weighted average

of roughly 40 estimates, which generally don't cover all moves into and out of a given CZ and which are allowed unrestricted heterogeneity. By exploding the number of parameters while discarding data, in other words, Chetty and Hendren's solution makes a painful tradeoff.

We address the dimensionality problem, without compromising on the core causal model, without adding parameters, and without discarding data. Counter-intuitively, we do so while increasing the analysis population by a factor of XX by expanding the number of cohorts using both historical and more-recent administrative tax records, expanding the coverage of moves with parents' location histories from birth for all cohorts from 1978-1998, and using the larger population to allow for less-common moves and sequences of moves; another benefit of improving the coverage of moves is that we reduce attenuation bias and increase the variance of the key explanatory variable. But how, then, do we make estimation feasible in this larger regression model? Instead of breaking the estimation into pairwise sub-problems, we break the estimation into a residualization step for the move-specific fixed effects, and an estimation step for the exposure effects, using the Yule-Frisch-Waugh-Lovell theorem and recent results on inference to show equivalence to the efficient, infeasible estimator.

These methodological contributions lead to an empirical contribution, increasing the precision of exposure effect estimates by two orders of magnitude. In turn, increased precision makes it possible to provide the first formal test of the estimator's identifying assumption, that family- and child-specific unobservables are unrelated to location exposures *in every place* after accounting for the places families have moved: we show that looking only at within-family variation in exposure times, estimates are significantly different in only five out of 761 CZs. Overall, we conclude that our exposure effect estimates are both precise and causal, meeting the basic criteria needed to rank and recommend relocations.

We use our exposure effect estimates to revisit several outstanding questions in the academic literature, and pose new questions for practice.

First, we ask whether observed outcomes, which are currently used for most rankings and policy recommendations, are biased relative to causal exposure effects: we measure bias using our estimates as a benchmark, and find that observed outcomes dramatically overstate exposure effects for major urban areas, the Upper Midwest, the Great Plains, and the Mountain states, while understating exposure effects for the South and Southwest. Thus, recommendations based on observed outcomes are badly distorted relative to causal estimates. Nonetheless, observed outcomes are more precise, and we ask whether efforts to predict and rank "good" CZs could be improved by relying in part on observed outcomes. Building MSE-minimizing shrinkage estimates of exposure effects as in Chetty and Hendren (2018b) and Armstrong, Kolesár, and Plagborg-Møller (2022), we find that the estimates load almost entirely on our exposure effect estimates, rejecting prediction based on the heavily-

biased observed outcomes. We find similar results for all CZ population sizes, suggesting that observed outcomes should be replaced with exposure effects in rankings for any CZ.

Second, we defend the premise of Chetty and Hendren (2018b) that causal exposure effect estimates can be ranked, from the critique of Mogstad et al. (2024) that they cannot be made precise enough even for pairwise comparison. We find that a majority of estimates are significant in comparisons with the national mean and with other CZs, at conventional levels relevant for recommendations about when to move conditional on moving between any given pair of CZs. Corrections for multiple testing find that most comparisons of any meaningful magnitude are significant; this is especially true for comparisons representing the most likely relocations for families in our population. Even by Mogstad et al.'s stringent criterion holding the chance of *any* mis-ranking between any pair of CZs to less than 5 percent, it is possible with our estimates to limit the confidence intervals for most of the top 50 CZs to fewer than half of the other ranks, and for some fewer than a tenth. Given that most families will only ever rank one CZ (their origin) against any other CZs (their potential destinations), and given that the quasi-experiment speaks primarily to the benefits of earlier or later moves within pairs of CZs, we conclude that the estimates are more than adequate to guide the intended recommendations.

Third, changing perspective, we ask whether parents are likely to accept recommendations based on estimates of exposure effects. We rank CZs by parents' revealed preference, using the matrix of families moving from one CZ to another, and a modification of Google's PageRank (Page, 1996) algorithm proposed by Sorkin (2018), to iteratively identify "good" CZs by inflows from other "good" CZs in a form of weighted voting. This revealed preference ranking uncovers a serious problem for policies encouraging moves toward CZs with positive exposure effects for children: exposure effects are strongly negatively correlated with parents' revealed preferences. For example, the most desirable CZ, New York, has among the most negative exposure effects: New York is ranked near the bottom for families at the 25th percentile of parent family income and at the 75th percentile. We confirm Chetty and Hendren's (2018b) finding that mean monthly rent is lower in places with more-positive exposure effects, but we also find that mean monthly rent is higher in places that parents prefer: that is, parents find it worth paying more, despite the fact that exposure effects predict worse outcomes for their children.

What, then, is the appeal of New York and CZs like it? We make a connection with the new literature on place premiums in earnings (Card, Rothstein, and Yi, 2025), showing that there *is* an opportunity cost to greater exposure effects: the cost is job opportunities for parents, with a strong negative gradient between earnings premia and exposure effects, and a strong positive gradient between earnings premia and parents' revealed preference. Thus,

we find a novel, highly relevant tradeoff between parents' interests and children's interests.[1]

Finally, we address the use of exposure effects to make recommendations to heterogeneous populations. We ask in particular whether children of immigrants, who are targeted by a different set of policies than children of natives, would benefit from different recommendations than children of natives. Making use of a novel imputation from publicly-available administrative records (Klopfer and Miller, 2024), we identify the native-born children of immigrant parents and estimate their exposure effects. Different from the findings in Aloni and Avivi (2025) in Israel, our rankings for natives and immigrants are positively correlated, suggesting the potential to use the full-population estimates to inform policies for sub-populations. Finding that the estimates for the immigrant sub-population are less precise than the full-population estimates, we build MSE-minimizing shrinkage estimates, which again load primarily on causal exposure effects. Ranking CZs, we find that an informative ranking is possible for the top 25 CZs, covering two-thirds of the immigrant population.

2 Data

We use data from full-population federal income tax records. To improve quasi-experimental estimates based on the timing of family moves across commuting zones (CZs), we expand the population in Chetty and Hendren (2018b) by a factor of XX, and we increase the variance in CZ exposures by measuring family residence from birth onwards. To study heterogeneity by parent nativity, we expand coverage of nativity in Chetty et al. (2020 and 2026) by a factor of 5, from a one-fifth sample to the full population. We also reduce uncertainty by sharpening parent and child income ranks with a longer panel of outcomes. The remainder of this section presents population and variable definitions, followed by summary statistics.

2.1 Population Definitions

Our base population consists of all children born from 1978-1998.

To rank children by income in childhood, we restrict the population to children with (1) a valid Social Security Number (SSN) issued by age 18 according to the sequence of issuance,[2]

[1] We caution, however, that a substantial body of causal evidence (and common sense) connects greater parental earnings with better outcomes for children; fully establishing the tradeoff between parents' and children's outcomes requires that income effects be quantified.

[2] Individuals with Individual Taxpayer Identification Numbers (ITINs) are non-citizens. SSNs may be issued to foreign-born citizens and non-citizens, though non-citizen status is recorded and rarely updated. Foreign-born non-citizens are not eligible to naturalize and obtain citizenship until they reach age 18, making it unlikely that a citizen-status SSN would be issued under age 18. An exception under the Child Citizenship Act of 2000 has been available since February 27, 2001, affecting children born after the same date in 1983, for foreign-born children adopted by citizen parents or living with naturalized parents. Our age cutoff does

and (2) no record of non-citizen status.[3] These conditions ensure that children were citizens at birth, with parents subject to taxation as citizens or residents.

To rank children by income in adulthood, we exclude children deceased as of 2025, and children not observed by the IRS from 2022-25.[4]

Parent and Child Linkages.—We define parents as the first tax filers to claim each child as a dependent between 1994 and 2016, such that female filers and single male filers are aged 15-40 at the child's birth. Relative to Chetty and Hendren (2018b) who study the 7 cohorts born from 1980-86, we study the 21 cohorts born from 1978-98, using recovered tax returns from 1994-95 (Alexander et al., 2024) to extend coverage to the birth cohorts of 1978-79.[5] We link parents to XX percent of children claimed through age 18.

Siblings.—We define siblings as children linked to the same tax filers. Studying a longer panel of birth cohorts increases the number of linked siblings more than proportionately, since nearly all siblings are found in different birth cohorts.

Stayers and Movers.—We define stayers as children whose parents lived one CZ, and movers as children whose parents lived in more than one CZ, through age 23.

We implement our definitions of stayers and movers using recovered tax returns from 1974-95 (Alexander et al., 2024) that allow us to observe parents' residence and identify moves from birth for nearly all children. We exclude the X.XX percent of children who have no valid residence. Measuring early moves identifies a larger population of movers, increases the variance of place exposures, and reduces attenuation bias in exposure estimates.

Second-Generation Natives, Mixed-Natives, and Immigrants.—We define second-generation natives as children with only native-born parents, second-generation mixed-natives as those with a native parent and a foreign-born parent, and second-generation immigrants as those with only foreign-born parents.[6] We shorthand these children as "native," "mixed-

not exclude natives: Klopfer and Miller (2024, Figure 4) show that native-born children from the 1978 birth cohort only rarely obtain an SSN after age 10, from 1983 after age 6, and from 1991 after age 1.

[3]Current citizenship status is recorded in the IRS DM-1 file and updates since 2014 are recorded in the DM-1-SSN file. Citizenship was not recorded in the DM-1 file until mid-1981, and a small number of children born and issued SSNs before mid-1981 do not have a recorded citizenship status. We include these SSNs, because most are held by citizens, and native birth is more likely the earlier the SSN is issued.

[4]Individuals are observed by the IRS if they are reported on any tax return (Form 1040) as a filer, spouse, or dependent; if they are reported on any information return (Form W-2, Form 1099, and others) or health coverage return (Form 1095); or if they are the subject of reporting from state and federal prison systems. Individuals not observed by the IRS are most often deceased or non-resident non-citizens.

[5]We do not use tax returns prior to 1994 for parent and child linkages because the 1969, 1974, 1979, 1984, and 1989 tax return files do not include dependent SSNs.

[6]In early papers based on IRS records alone, Chetty et al. (2014) and Chetty and Hendren (2018a and b) did not have access to parent nativity. Our definition differs from that used at the Census Bureau by Chetty et al. (2020 and 2026) in two important respects. First, Chetty et al. include first-generation immigrant children who moved to the United States with their parents: these children are not citizens by birth, and are more likely to emigrate. Second, Chetty et al. define immigrant children by mothers' nativity: mixed-

native," and "immigrant" throughout the paper, though all are native-born.

We implement our definitions of second-generation nativity by probability weighting as in Klopfer and Miller (2024): this gives us population coverage with little loss of accuracy relative to survey linkages.[7] Weights for children are single parents' individual weights, and the product of married parents' individual weights; we define parents' weights (probabilities of native or foreign birth) in Section 2.2. For example, if the probability that the first parent is foreign-born is 0.9, and the probability that the second is foreign-born is 0.2, the probability that both are foreign-born is 0.18. The weighted populations are primarily attributable to high-certainty observations. Figure **??** displays the probabilities that a child has only native, mixed-native, or only immigrant parents on the probability simplex: there is little overlap between the native and immigrant populations, and little imputation uncertainty.[8]

2.2 Variable Definitions

We define parent and child variables in the tax data, starting with parent geography, nativity, and income ranks before discussing child income ranks. All monetary values are reported in 2012 dollars.

Parent Geography.—In each year, we define parent ZIP codes using the first valid ZIP code found, in order, on the mother's tax return, the father's tax return, the plurality of the mother's information returns, or the plurality of the father's information returns. We translate ZIP codes to 1990 CZs using a crosswalk covering ZIP codes valid in 1974-2022, adding coverage for Puerto Rico using 2020 CZs and for overseas territories by postal abbreviation. We describe how we built the crosswalk in Online Appendix **??**.

When defining moves, ages at moves, and years of exposure to each CZ, we treat the CZ observed closest to birth (or the later if two are equally close) as the first location. Each time a child is observed in a different CZ, we treat that CZ as the next location. To account for missing tax return data between 1974, 1979, 1984, 1989, and 1994, and for non-filing, we define the year of each move as the midpoint between the first year observed at the new location and the last year observed at the last location, rounded up. We define the age at

nativity parents differ from foreign-nativity parents because of selection, because they more commonly obtain citizenship by marriage, and because they cannot be single parents. With population-scale data, and given evidence that these children's outcomes differ (Ramakrishnan, 2004), we see value in making the distinction. We leave a study of first-generation immigrant mobility to future work.

[7]Weights from SSN application records are often as accurate as the survey responses Chetty et al. (2020 and 2026) rely on: in an extract based on their sample definition, 5.1 percent of weighted birthplace responses in the 2000 decennial census long form and 2005-15 ACS were imputed.

[8]Probability weighting (fractional imputation) improves accuracy over classification based on the same underlying data (as in Yonker, 2017; Amornsiripanitch et al., 2023; Bernstein et al., 2025; and Cookson, Guttman-Kenney, and Mullins, 2025), see Kim and Fuller (2004) and Kim and Rao (2009).

move as the year minus the child's birth year, and consider moves from age 1 to 23. Because tax returns and information returns are filed after the end of the corresponding calendar year, we calculate exposures for the first location from age 0 to the first move, for each following location from the age a child moved in and out, and for the final location from the last move to age 24. For example, consider a child born in 1995 whose parents file 1995 taxes from one CZ, and file 1996-2019 taxes from another CZ: we would code that child as having lived in the first CZ for 1 year, and the second CZ for 23 years.[9]

Parent Nativity.—We define parent nativity by probability weighting as in Klopfer and Miller (2024). We link parents by the first five digits of their SSNs to Klopfer and Miller's crosswalk for the SSN issue month and state. We obtain each parent's sex and birth year from the IRS DM-1 file, define the issue year from the issue month, and define the issue age from the issue year and birth year. Finally, we link parents by sex and the SSN issue age, year, and state to Klopfer and Miller's probability table, which is itself derived from and validated against the original SSN application records.

For children born to parents aged 30 in 1978-1998 who were issued SSNs through age 30, Klopfer and Miller (2024, Figure 4 and 5) show that the MSE of the imputed probabilities relative to the true probabilities is typically on the order of 0.033, and at most 0.166.

Parent Family Income and Percentile Rank.—We follow Chetty et al. (2014) and subsequent papers in defining each parent's family income as the total pre-tax income by tax unit. In years when a parent files a tax return (Form 1040) as a primary or secondary filer, we define that parent's family income as the sum of Adjusted Gross Income, tax-exempt interest, and non-taxable Social Security and Disability (SSDI) benefits reported on the tax return. When no parent files, we define each parent's family income as the sum of wage earnings (from Form W-2), unemployment benefits (from Form 1099-G), and taxable and nontaxable Social Security and Disability benefits (from Form SSA-1099) for both parents. When no parent is observed, we define family income as zero.[10]

To measure parent family income with children as the unit of observation, we average family income across both parents regardless of their marital status. For parents who file

[9]To see why this simple calculation is roughly correct, suppose the child was born on July 1, 1995, and his parents filed taxes on April 15, 1996 and moved on July 1, 1996: in that case, the child would have lived six months of 1995 in the first CZ by the time taxes are filed, and another six months of 1996 in the first CZ before moving to the second; given that the median child is born in the summer, most filers file on tax day, and most children move during the summer break, this is a fair approximation. Further, if the parents always file and move later, the child will spend at most a few months more in the first CZ, and the same few months less in the last CZ, but for intermediate moves, the errors balance.

[10]Chetty et al. (2014) note that parents who are not observed have no taxable income. However, non-resident non-citizen parents may have no taxable income but significant non-taxable income (Klopfer and Miller, 2024) which would distort parent ranks for immigrant children. We focus our analysis on children who were citizens at birth to ensure that parents were subject to taxation as citizens or residents.

together, average family income equals their individual family incomes obtained from their joint tax return. If the parents do not file together, average family income equals the average of their new tax units' family incomes, and may be reduced by one or both of the parents being in a single-income tax unit.

To standardize the measurement of parents' income across cohorts, and to reduce the influence of annual variation in parents' family income, we average parent family income across all years in which the child is aged 0-23; we over-weight income from tax returns in 1979, 1984, 1989, and 1994 to account for tax years 1978-81, 1982-86, 1987-91, and 1992-94. This definition covers almost 5 times the age span available to Chetty and Hendren (2018b), and better measures family resources during early and middle childhood.

We exclude the X.XX percent of children whose parents have zero or negative family income after averaging, since parents filing with zero or negative net income generally have positive gross income and would be mis-ranked.

We define parent family income ranks within each full birth cohort. This ensures that ranks do not vary systematically with different years of tax records for different cohorts.

Child Family Income and Percentile Rank.—We define child family income identically to parent family income. Unlike Chetty et al. (2014) and related papers, we average child family income across all available ages from 24 onward, with the final year of income measurement in 2022. We do this to reduce the influence of annual variation in children's family income: for example, for 1978, we aggregate 21 years of child family income from 2002-2022 in place of the single year ranked by Chetty and Hendren (2018b). We allow for zero or negative incomes. However, zero incomes are uncommon due to the requirements that children be alive, present in the tax data (which include information returns in all outcome years), and citizens (and therefore taxed on income globally).

We define child family income ranks within each full birth cohort. Similarly to the ranking for parent family income, this ensures that ranks do not vary systematically with different age ranges and different years of tax records for different cohorts.

2.3 Summary Statistics

Table **??** reports summary statistics for the full population and weighted second-generation native, mixed-native, and immigrant populations. Panel A reports statistics for the full population, Panel B reports statistics for stayers, and Panel C reports statistics for movers who moved up to 6 times (XX percent of all movers).

The primary populations for analysis are the total and weighted movers populations: the total and weighted counts of native, mixed-native, and immigrant movers are XX.X, XX.X,

X.X, and X.X million, respectively. Our population of between-CZ movers is XX times larger than the population in Chetty and Hendren (2018b, Table 1). Our coverage of children with parents' nativity is 5 times larger than in Chetty et al. (2026, Table 1). Taken together, our weighted native, mixed-native, and immigrant movers populations are XX, X, and X times the total population of between-CZ movers in Chetty and Hendren (2018b, Table 1). Larger populations improve the precision of movers estimates overall and for sub-populations.

3 Methods

We begin by defining our estimands, exposure effects, for each CZ. We next define our estimators, and the assumptions required to identify exposure effects. For each estimator, we discuss estimation in large populations with many nuisance parameters. Finally, we discuss variance-covariance estimators, assumptions, and estimation.

3.1 Estimands

We seek to understand how childhood exposures to locations $c \in C$ at ages $a \in A$, captured by location histories $\boldsymbol{c}_i = (c_{ia} : a \in A)'$, affect children's outcomes in adulthood. Specifically, we seek to estimate parameter vector $\boldsymbol{\mu}$ in the statistical model

$$y_i = (\boldsymbol{\mu}^0 + \boldsymbol{\mu}^1 p_i) \cdot \boldsymbol{e}_i + (\boldsymbol{\kappa}^0 + \boldsymbol{\kappa}^1 p_i) \cdot \boldsymbol{s}_i + \theta_i \tag{1}$$

where y_i is a child's family income rank or other outcome; p_i is the parents' family income rank; $\boldsymbol{\mu} = (\boldsymbol{\mu}^{0\prime}, \boldsymbol{\mu}^{1\prime})'$ summarizes causal effects of location exposures $\boldsymbol{e}_i = (\sum_A 1_c(c_{ia}) : c \in C)'$; $\boldsymbol{\kappa} = (\boldsymbol{\kappa}^{0\prime}, \boldsymbol{\kappa}^{1\prime})'$ summarizes causal effects of location sequences $\boldsymbol{s}_i = (1_{\boldsymbol{d}}(\boldsymbol{d}_i) : \boldsymbol{d} \in \boldsymbol{D})'$ where we define $\boldsymbol{d}_i = (c_{ia} : c_{ia} \neq c_{ia-1}, a \in A)'$ as the vector composed of the origin and each move's destination; and θ_i summarizes other family- and child-specific factors.

This simplified model restricts location-effect heterogeneity. First, we assume exposure and sequence effects are additively separable, ruling out order effects, complementarity, and substitutability. Second, we assume linear interactions between parental income ranks and exposure and sequence effects. Third, we assume constant (and therefore age-independent) exposure effects during childhood and zero exposure effects otherwise.[11] Fourth, we assume constant sequence effects during childhood and zero sequence effects otherwise.

[11] We rely on past validations of the second and third assumptions. Chetty et al. (2014) show linearity in parental income ranks. Chetty and Hendren (2018a) and Chetty et al. (2026) show constant exposure effects through early adulthood and zero effects thereafter.

3.2 Estimators and Assumptions

Ideally, we would estimate location effects by randomly assigning families to location histories, as in the Moving to Opportunity experiment (see Chetty, Hendren, and Katz, 2016). However, our goal is to estimate effects for all potential moves and family income ranks, which is more readily done by observing the population of voluntary movers in tax data.

To estimate exposure effects $\boldsymbol{\mu}$, we rely on the timing of parents' moves across locations to obtain variation in children's location exposures, following Chetty and Hendren (2018b). Consider two children who have different location exposures but the same family income rank, location sequence, and other family inputs. In this case, we can isolate the exposure effects at the children's shared rank p; if we observe comparable children at multiple ranks, then we can also distinguish the intercept and slope of the exposure effects. For example, suppose the children move from New York to Los Angeles, at different ages. The child who moves at a younger age has a longer exposure in Los Angeles, and a shorter exposure in New York, so comparing outcomes identifies the rank-specific exposure effect

$$(y_i - y_j)/(e_{i,NY} - e_{j,NY}) = (\mu^0_{NY} + p\mu^1_{NY})$$

for New York relative to Los Angeles, normalizing the exposure effect of Los Angeles to zero. That is, we can identify exposure effects if we can make other family inputs comparable.

Estimator A1. A natural estimator of $\boldsymbol{\mu}$ is the sequence fixed effects specification

$$y_i = (\boldsymbol{\mu}^0 + \boldsymbol{\mu}^1 p_i) \cdot \boldsymbol{e}_i + (\boldsymbol{\alpha}^0 + \boldsymbol{\alpha}^1 p_i) \cdot \boldsymbol{s}_i + \epsilon_i \tag{2}$$

where $(\boldsymbol{\alpha}^0 + \boldsymbol{\alpha}^1 p_i) \cdot \boldsymbol{s}_i = (\boldsymbol{\kappa}^0 + \boldsymbol{\kappa}^1 p_i) \cdot \boldsymbol{s}_i + \bar{\theta}_{p_i \boldsymbol{s}_i}$ incorporates both causal effects $\boldsymbol{\kappa}$ and selection effects $\bar{\theta}_{p_i \boldsymbol{s}_i} = \mathrm{E}[\theta_i | \boldsymbol{s}_i, p_i \boldsymbol{s}_i]$, and ϵ_i captures residual variation in other family inputs. This estimator absorbs sequence effects, and identifies exposure effects from variation due to the timing of moves; similar to the estimator in Aloni and Avivi (2025), it extends the estimator in Chetty and Hendren (2018b) by allowing for multiple-move sequences.

To identify $\boldsymbol{\mu}$, the error must have a mean of zero with respect to regressors $\boldsymbol{e}_i$ and $p_i \boldsymbol{e}_i$, after absorbing the fixed effects:

Assumption A1. Conditional on location sequences and their interactions with parent family income rank $\boldsymbol{s}_i$ and $p_i \boldsymbol{s}_i$, exposures and their interactions with parent family income rank $\boldsymbol{e}_i$ and $p_i \boldsymbol{e}_i$ are jointly orthogonal to residual variation in childrens' outcomes

$$\mathrm{E}[\epsilon_i | \boldsymbol{e}_i, p_i \boldsymbol{e}_i, \boldsymbol{s}_i, p_i \boldsymbol{s}_i] = \mathrm{E}[\epsilon_i | \boldsymbol{s}_i, p_i \boldsymbol{s}_i]$$

Intuitively, Assumption A1 states that exposures cannot be related to outcomes through selection, after accounting for children's location sequences. Assumption A1 also states that the interaction of parent ranks and exposures cannot be related to outcomes through selection, after accounting for the interaction of parent ranks and children's location sequences. For example, conditional on moving to Los Angeles from New York, families should not move later when their unobservables are better. Further, it shouldn't be the case that higher-rank parents move later and lower-rank parents move earlier when their unobservables are better. That is, there should be no selection on average or by parent ranks.

We test Assumption A1 in Section 4.3, extending and formalizing the specification tests in Chetty and Hendren (2018a and 2018b) to find stronger evidence against either systematic or location-specific bias in estimates from Equation A1.

3.3 Estimation

We exploit the Yule-Frisch-Waugh-Lovell theorem to jointly estimate the 1,522 CZ intercept and slope parameters $\boldsymbol{\mu}$[12] as

$$\hat{\boldsymbol{\mu}} = (\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}})^{-1}\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{Y}} \tag{3}$$

after we obtain the residuals of childrens' outcomes and location exposures, y_i and $\boldsymbol{x}_i = (\boldsymbol{e}_i', p_i\boldsymbol{e}_i')'$ on location sequences $\boldsymbol{z}_i = (\boldsymbol{s}_i', p_i\boldsymbol{s}_i')'$.

The key technical challenge for estimation is how to obtain the residuals. Direct residualization on $\boldsymbol{Z}$ is infeasible: storing and inverting $\boldsymbol{Z}'\boldsymbol{Z}$ would require tens of terabytes of memory. Iterative residualization is inefficient, requiring days of server time and weeks for specifications that include additional slope terms for each location sequence.[13]

We develop an efficient residualization method, relying on the sparsity of $\boldsymbol{Z}$ and $\boldsymbol{X}$: first, $\boldsymbol{z}_{i\boldsymbol{d}} = (1_{\boldsymbol{d}}(\boldsymbol{d}_i), p_i 1_{\boldsymbol{d}}(\boldsymbol{d}_i))' = \boldsymbol{0}$ for $\boldsymbol{d} \neq \boldsymbol{d}_i$; second, $\boldsymbol{x}_{ic} = (e_{ic}, p_i e_{ic})' = \boldsymbol{0}$ for $c \notin \boldsymbol{d}_i$. The

[12] We observe exposures for 761 CZs: the 741 1990 CZs covering the states and the District of Columbia, 10 2020 CZs covering Puerto Rico, and 10 pseudo-CZs defined by postal abbreviation for overseas territories. We normalize $\mu^0_{LA} = \mu^1_{LA} = 0$ due to collinearity, treating Los Angeles as the base category.

Joint estimation of both $\boldsymbol{\mu}$ and $\boldsymbol{\kappa}$ by OLS is infeasible, with millions of intercept and slope parameters across all location sequences. The approach in Chetty and Hendren (2018b), two-step estimation of $\boldsymbol{\mu}$ by OLS within location sequences, is feasible but inefficient: although their second step constrains location exposure parameters to be equal across all location sequences, their first step allows for heterogeneity in exposure parameters across location sequences; consequently, they are forced to estimate orders of magnitude more parameters than necessary, and to discard many observations for which the parameters are not identified.

[13] In alternative specifications, we have also included interactions of location sequences with birth cohorts and squared birth cohorts, as in Chetty and Hendren (2018b). However, the primary reason to include cohort interactions is to address differential measurement error and selection across cohorts (Chetty and Hendren 2018a, Online Appendix A). Doing so is not necessary when we measure location histories from birth and, because it discards up to a tenth of the population due to small cell sizes for rare location sequences, doing so is inefficient. Further, the correlation between estimates is near one.

sparsity of $\boldsymbol{Z}$ allows us to residualize $\boldsymbol{Y}$ and $\boldsymbol{X}$ separately for each sub-population

$$\tilde{\boldsymbol{Y}} = (\tilde{\boldsymbol{Y}}_{\boldsymbol{d}}' : \boldsymbol{d} \in \boldsymbol{D})'$$
$$\tilde{\boldsymbol{X}} = (\tilde{\boldsymbol{X}}_{\boldsymbol{d}}' : \boldsymbol{d} \in \boldsymbol{D})'$$

where we compute residuals for each sub-population, by multiplication with the annihilator matrix $\boldsymbol{M_{dd}} = \boldsymbol{I} - \boldsymbol{Z_{dd}}(\boldsymbol{Z_{dd}}'\boldsymbol{Z_{dd}})^{-1}\boldsymbol{Z_{dd}}'$, as $\tilde{\boldsymbol{Y}}_{\boldsymbol{d}} = \boldsymbol{M_{dd}}\boldsymbol{Y_d}$ and $\tilde{\boldsymbol{X}}_{\boldsymbol{d}} = \boldsymbol{M_{dd}}\boldsymbol{X_d}$. The sparsity of $\boldsymbol{X}$ allows us to compute $\tilde{\boldsymbol{X}}_{\boldsymbol{d}}$ as $\tilde{\boldsymbol{X}}_{\boldsymbol{dc}} = \boldsymbol{M_{dd}}\boldsymbol{X_{dc}}$ for $c \in \boldsymbol{d}_i$ and $\tilde{\boldsymbol{X}}_{\boldsymbol{dc}} = \boldsymbol{0}$ for $c \notin \boldsymbol{d}_i$. Our method, implemented in Stata as `residby`, runs 3-4 orders of magnitude faster than the current benchmark absorbing categorical-by-continuous interactions, `hdfe` (Correia, 2016);[14] the performance gap is larger for alternative specifications with additional slope interactions. Since many problems share this structure, our method may be of wider interest.

3.4 Variance and Covariance

Given OLS estimates $\hat{\boldsymbol{\mu}}$, we estimate the variance-covariance matrix (VCE) to rank locations by their exposure effects.

We implement the VCE estimator of Abadie et al. (2020), following their design-based approach to inference as suggested by Mogstad et al. (2024). The design-based approach is motivated by the fact that there is no sampling uncertainty in population data. The source of uncertainty is, instead, that exposures could have taken different values if children had moved at different ages. In the nomenclature of Abadie et al. (2020), we treat location exposures as causal and location sequences as descriptive.

The Abadie et al. (2020) (AAIW) VCE estimator is

$$\hat{\boldsymbol{V}}_{\hat{\boldsymbol{\mu}}}^{AAIW} = (\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}})^{-1}\tilde{\boldsymbol{U}}_i'\tilde{\boldsymbol{U}}_i(\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}})^{-1} \tag{4}$$

where $\boldsymbol{u}_i = \tilde{\boldsymbol{x}}_i\hat{\epsilon}_i$, and $\hat{\epsilon}_i = \tilde{y}_i - \tilde{\boldsymbol{x}}_i'\hat{\boldsymbol{\mu}}$. We residualize $\boldsymbol{U}$ on $\boldsymbol{Z}$ separately for each sub-population

$$\tilde{\boldsymbol{U}} = (\tilde{\boldsymbol{U}}_{\boldsymbol{d}}' : \boldsymbol{d} \in \boldsymbol{D})'$$

and as before compute $\tilde{\boldsymbol{U}}_{\boldsymbol{d}}$ as $\tilde{\boldsymbol{U}}_{\boldsymbol{dc}} = \boldsymbol{M_{dd}}\boldsymbol{U_{dc}}$ for $c \in \boldsymbol{d}_i$ and $\tilde{\boldsymbol{U}}_{\boldsymbol{dc}} = \boldsymbol{0}$ for $c \notin \boldsymbol{d}_i$. Intuitively, the AAIW VCE estimator removes "uncertainty" due to estimated exposure parameter heterogeneity across location sequences.[15]

[14] We obtain these results in part by performing only the necessary computations, and in part by coding the computations efficiently: we build on the compiled-loop implementation of `regressby` (Stepner, Droste, and Townsend, 2018; Caceres-Bravo, 2018), which we modify to parallelize residualization.

[15] For comparison, the Eicker-Huber-White VCE estimator is $\hat{\boldsymbol{V}}_{\hat{\boldsymbol{\mu}}}^{EHW} = (\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}})^{-1}\boldsymbol{U}'\boldsymbol{U}(\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}})^{-1}$ where $\boldsymbol{U}$

To enable secondary analyses, we will release AAIW and EHW VCE estimates: as Chetty and Hendren (2018b), Mogstad et al. (2024), and Kline (2025) observe, covariance terms are generically nonzero, so that the VCE cannot be summarized by standard errors.

4 Exposure Effect Estimates

This paper's primary objective is to estimate location exposure effects that can be relied on as an input to decisions by families and policymakers. This section reports estimated exposure effects, assesses the degree to which they reflect signal or noise, and tests the identification assumption required to interpret them as causal.

4.1 Baseline Estimates

What is the impact of location exposures on children's family income ranks in adulthood, given their parents' family income ranks?

We measure CZ exposure effects relative to the national mean, and their uncertainty, by combining estimates from equations (3) and (4),

$$\hat{\boldsymbol{\mu}}_p = \boldsymbol{Q}\boldsymbol{A}_p\hat{\boldsymbol{\mu}} \tag{5}$$

$$\hat{\boldsymbol{V}}_{\hat{\boldsymbol{\mu}}_p} = \boldsymbol{Q}\boldsymbol{A}_p\hat{\boldsymbol{V}}_{\hat{\boldsymbol{\mu}}}\boldsymbol{A}_p'\boldsymbol{Q}' \tag{6}$$

where $\boldsymbol{A}_p = [\boldsymbol{I}\ p\boldsymbol{I}]$ adds intercepts to slopes multiplied by parent ranks and $\boldsymbol{Q} = \boldsymbol{I} - \mathbf{1}\boldsymbol{\pi}'$ subtracts the exposure-weighted mean based on exposure shares $\boldsymbol{\pi} = \bar{\boldsymbol{e}}/\sum_C \bar{\boldsymbol{e}}_c$.

For example, children with parents at the 25th percentile of parent family income, exposed to Los Angeles for an additional year of childhood relative to the national mean, have child family income $\hat{\mu}_{LA25} = -X.XXXX$ ($X.XXXX$) percentiles lower. At the 75th percentile of parent family income, exposure to Los Angeles for an additional year results in child family income $\hat{\mu}_{LA75} = -X.XXXX$ ($X.XXXX$) percentiles lower. To obtain the full effects of location exposures, we scale effects by 24 years of exposure, as in the estimation.[16] For example, at the 25th percentile, a full childhood of exposure to Los Angeles lowers children's family incomes by $\hat{\mu}_{LA25} \times 24 = -X.XXX$ ($X.XXX$) percentiles; at the 75th percentile, full exposure lowers children's family incomes by $\hat{\mu}_{LA75} \times 24 = -X.XXX$ ($X.XXX$) percentiles.

has not been residualized. In practice, the VCE estimates are similar: AAIW standard errors for exposure effects are roughly X percent smaller than EHW standard errors. These findings support the assumption in Section 3.1 that exposure and sequence effects are additively separable.

[16] We measure childhood locations from age 0 through 23, for a total of 24 years of exposure, based on the evidence in Chetty and Hendren (2018a) and Chetty et al. (2026) that exposure effects are approximately linear from age 3 through 23.

Figure **??** and **??** report exposure effects $\hat{\mu}_{c25}$ and $\hat{\mu}_{c75}$, and their confidence intervals, for the 50 largest CZs and for all CZs.

How large are exposure effects relative to the national mean? One relevant comparison is with parent income effects on child outcomes. We equate location exposure effects and parent income effects using the slope of child income ranks relative to parent income ranks among stayers, $X.XXXX$. Across all CZs the standard deviation of full exposure effects weighted by population, $X.XXXX \times 24 = X.XXX$ for the 25th percentile of parent income or $X.XXXX \times 24 = X.XXX$ for the 75th percentile of parent income,[17] is equivalent to having parents $X.XXX/X.XXXX = X.XXX$ or $X.XXX/X.XXXX = X.XXX$ percentiles higher or lower in the income distribution. This evidence suggests that exposure effects are important for children's outcomes.

How significant are exposure effects relative to the national mean? Among the 50 largest CZs, at the 25th and 75th percentiles of parent income, $XX/50$ and $XX/50$ estimated exposure effects have a t-statistic greater than 1.960; accounting for multiple testing with a (conservative) Bonferroni correction, $XX/50$ and $XX/50$ have a t-statistic greater than 3.292. Among all CZs, at the 25th and 75th percentiles of parent income, $XXX/761$ and $XXX/761$ have a t-statistic greater than 1.960; accounting for multiple testing with a Bonferroni correction, $XXX/761$ and $XXX/761$ have a t-statistic greater than 3.993.[18] Since many estimates are close to zero, not all estimates will be significant in tests against the national mean.

Figure **??** and **??** map the geography of exposure effects $\hat{\mu}_{c25}$ and $\hat{\mu}_{c75}$. Intergenerational mobility is higher for children who have more exposure to CZs in the Great Plains, the Ozarks, Appalachia, and the Mountain West; and less exposure to CZs in the South Atlantic, Gulf, and Pacific Coasts, and the Central Great Lakes. Major urban areas, notably Los Angeles, New York, and Chicago, under-perform the national mean and most rural areas. Geographic patterns are similar for the 25th and 75th percentile of parent incomes. Estimates supporting these observations are generally significant after Bonferroni correction.

4.2 Validating the Estimates: Signal and Noise

Imprecision is a source of risk in decisions based on estimates. Chetty and Hendren (2018b) raise concerns about the precision of exposure effect estimates, echoed in Mogstad et al.

[17] As we discuss in Section 4.2, the standard deviation of the estimated exposure effects is within X percent of the standard deviation of the true exposure effects.

[18] For X CZs at the 25th percentile of parent income, and X CZs at the 75th percentile of parent income, we treat estimates as insignificant regardless of their t-statistics because their confidence intervals for the effect of full exposure include values more than 50 percentile ranks away from the national mean.

(2024) and Andrews, Kitagawa, and McCloskey (2024). We now present a formal analysis showing that our methods fully address these precision concerns.

To understand the variation of the exposure effect estimates around the national mean, we would like to partition the total variance $\sigma^2_{\hat{\boldsymbol{\mu}}_p}$ into signal $\sigma^2_{\boldsymbol{\mu}_p}$ (true variance) and noise $\sigma^2_{\hat{\boldsymbol{\mu}}_p-\boldsymbol{\mu}_p}$ (variance due to estimation), with

$$\sigma^2_{\hat{\boldsymbol{\mu}}_p} = \sigma^2_{\boldsymbol{\mu}_p} + \sigma^2_{\hat{\boldsymbol{\mu}}_p-\boldsymbol{\mu}_p}$$

We estimate these unobservable variances from demeaned estimates $\hat{\boldsymbol{\mu}}_p$ and their variances $\hat{\boldsymbol{V}}_{\hat{\boldsymbol{\mu}}_p}$ from equations (5) and (6), as

$$\hat{\sigma}^2_{\hat{\boldsymbol{\mu}}_p} = \hat{\boldsymbol{\mu}}_p'\boldsymbol{\Pi}\hat{\boldsymbol{\mu}}_p \tag{7}$$

$$\hat{\sigma}^2_{\hat{\boldsymbol{\mu}}_p-\boldsymbol{\mu}_p} = \mathrm{tr}(\boldsymbol{\Pi}\hat{\boldsymbol{V}}_{\hat{\boldsymbol{\mu}}_p}) \tag{8}$$

$$\hat{\sigma}^2_{\boldsymbol{\mu}_p} = \hat{\sigma}^2_{\hat{\boldsymbol{\mu}}_p} - \hat{\sigma}^2_{\hat{\boldsymbol{\mu}}_p-\boldsymbol{\mu}_p} \tag{9}$$

where we weight the estimates and their variances by exposure shares $\boldsymbol{\Pi} = \mathrm{diag}(\boldsymbol{\pi})$.[19] One way to summarize precision based on these estimates is the signal to noise ratio, $\hat{\sigma}^2_{\boldsymbol{\mu}_p}/\hat{\sigma}^2_{\hat{\boldsymbol{\mu}}_p-\boldsymbol{\mu}_p}$.

Table **??** reports the variance components. We report variances across all CZs, for the 25th and 75th percentiles of parent family income, in Panel A, Columns 1 and 2. The signal variances for exposure effects at the 25th and 75th percentiles are $X.XXXX$ and $X.XXXX$, or XX and XX percent of the total variances; the signal standard deviations are $X.XXX$ and $X.XXX$ percentile ranks per year of exposure; or XX and XX percent of the total standard deviations. In other words, the noise variances of $X.XXXX$ and $X.XXXX$ account for only X and X percent of the total variance, and the noise standard deviations of $X.XXXX$ and $X.XXXX$ are XX and XX percent of the total standard deviations. Summarizing the variance components, we find signal to noise ratios of XX.XX and X.XX at the 25th and 75th percentiles, both excellent.

We build on the contribution of Chetty and Hendren (2018b) by improving the precision of their estimates. In Table **??**, Panel B, Columns 1 and 2, we report estimates using the replication data from Chetty and Hendren (2018b), weighting by population for comparability because this is the weight they used to demean their estimates. This analysis reveals two new findings. First, compared to their estimated noise variances of X.XXXX and X.XXXX, our estimates reduce noise by a factor of XX and XX, respectively. Second, compared to their estimated signal variances of X.XXXX and X.XXXX, our estimates suggest that the

[19] By a standard result for quadratic forms, if $\mathrm{E}[\hat{\boldsymbol{\mu}}_p] = \boldsymbol{\mu}_p$ and $\mathrm{Var}[\hat{\boldsymbol{\mu}}_p] = \boldsymbol{V}_{\hat{\boldsymbol{\mu}}_p}$, then $\mathrm{E}[\hat{\boldsymbol{\mu}}_p'\boldsymbol{\Pi}\hat{\boldsymbol{\mu}}_p] = \boldsymbol{\mu}_p'\boldsymbol{\Pi}\boldsymbol{\mu}_p + \mathrm{tr}(\boldsymbol{\Pi}\boldsymbol{V}_{\hat{\boldsymbol{\mu}}_p})$. With weights $\boldsymbol{\pi}$ and $\hat{\boldsymbol{\mu}}_p'\boldsymbol{\pi} = 0$, the quadratic forms are variances.

true signal variances are approximately XX and XX percent smaller. Re-estimating the noise in the original estimates by subtracting our signal variance estimates from the original total variance estimates, we find that the original signal to noise ratios may have been X.XX and X.XX or lower, much lower than the ratios of 0.51 and 0.25 implied by the original variance estimates. Taken together, these findings suggest signal to noise ratios XX and XX times greater for our estimates than for the original estimates, or two orders of magnitude greater precision.

Are exposure effect estimates more informative for larger CZs? Table **??**, Columns 3 to 10 report variance components by CZ population, starting with the first 50, next 150, next 395, and final 166 CZs by population. The first three categories cover all CZs considered by Chetty and Hendren (2018b), and the fourth category covers all CZs added in our analysis. We find that estimates are most informative for the 50-largest CZs, and among high-income families in particular, with signal to noise ratios of XX and XXX at the 25th and 75th percentiles, compared with 1.27 and 2.55 for the original estimates. For the remaining categories, signal to noise ratios are flat at XX for the 25th percentile, while signal to noise ratios decline gradually at the 75th percentile, from X for CZs 51-200, to X for CZs 201-595, to X for CZs 596-761. This pattern, of greater signal at the 75th percentile for CZs 1-50, and lower signal at the 75th percentile for all remaining CZs, is consistent with the concentration of 75th percentile families in major urban CZs, and fewer high-income families in smaller CZs.

The signal to noise ratio for every population category of our estimates is substantially greater than the signal to noise ratio even for the 50-largest CZs in the original estimates. Comparing noise variance estimates directly, it is only in our smallest, newly-added CZs that the noise variance exceeds the noise variance of the original estimates for the 50-largest CZs. Thus, estimates for most CZs appear to be highly informative.

Overall, our new estimates address the primary limitation of the estimates in Chetty and Hendren (2018b): precision. Our estimates provide an empirical answer to the critiques in Mogstad et al. (2024) and Andrews, Kitagawa, and McCloskey (2024). Sections 5.1, 5.2, 5.3, and 5.4 make use of our estimates to address challenges raised by these authors: bias assessment for non-causal observational estimates, forecasts of causal effects, pairwise comparisons between locations, and multiple comparisons among locations in the form of rankings.

4.3 Validating the Estimates: Identification

Bias is another source of risk in decisions based on estimates. The estimates $\hat{\boldsymbol{\mu}}_p$ from equation (5) are unbiased for the true causal parameters $\boldsymbol{\mu}_p$ if Assumption A1 is valid. The assumption is valid if family- and child-specific factors θ_i are orthogonal to *all* location exposures at *all* income levels $\boldsymbol{e}_i$ and $p_i\boldsymbol{e}_i$ after accounting for location sequences and their interactions with parent family income. That is, parents should not choose the duration of a given exposure based on unobservables, conditional on having a particular sequence of exposures.

We validate the estimated exposure effects as unbiased *for each location*, first by testing for bias due to family-level factors, and second by testing for bias due to child-level factors.

4.3.1 Family-Level Factors

Because families move together, selective exposures are less of a concern for within-family comparisons than for between-family comparisons. Based on the observation that older and younger siblings have different exposures for the same family moves, Chetty and Hendren (2018a, Section 5), Chetty et al. (2026, Appendix J), and Aloni and Avivi (2025, Appendix D) test for bias *on average in high- or low-mobility locations based on non-causal estimates of mobility.*[20] These analyses do not test Assumption A1, but instead a weaker assumption that stayers' outcomes do not predict selection on movers' exposures.

With coverage of siblings across many cohorts, we use family fixed effects to test Assumption A1, modifying Estimator A1 for comparison with the sibling fixed effects Estimator A2:

Estimator A2. An estimator of $\boldsymbol{\mu}$ eliminating the influence of family-level unobservables is the nested family fixed effects specification

$$y_i = (\boldsymbol{\mu}^0 + \boldsymbol{\mu}^1 \bar{p}_{f_i \boldsymbol{d}_i}) \cdot \boldsymbol{e}_i + \boldsymbol{\beta}^0 \cdot \boldsymbol{h}_i + \boldsymbol{\beta}^1 \bar{p}_{f_i \boldsymbol{d}_i} \cdot \boldsymbol{s}_i + \tilde{\epsilon}_i \tag{10}$$

where we replace sequence intercepts $\boldsymbol{s}_i$ with nested family-by-sequence intercepts $\boldsymbol{h}_i = (1_{f\boldsymbol{d}}(f_i, \boldsymbol{d}_i) : (f, \boldsymbol{d}) \in \boldsymbol{H})'$.[21] In both Estimator A2 and A1, we replace p_i with $\bar{p}_{f_i \boldsymbol{d}_i}$, the average parent rank over all siblings with the same location sequence, to address within-family estimation noise;[22] since p_i or $\bar{p}_{f_i \boldsymbol{d}_i}$ is included in the model as a source of heterogeneity,

[20] Specifically, these authors ask whether the rate at which movers' outcomes converge to stayers outcomes, $\bar{\boldsymbol{y}}_p$, changes when they include family fixed effects.

[21] Although $\boldsymbol{h}_i$ captures most variation in parent ranks, we continue to estimate sequence-by-parent-rank slopes $\bar{p}_{f_i \boldsymbol{d}_i} \boldsymbol{s}_i$ for comparability with Estimator A1.

[22] Because parent family income ranks are similar within families, but child family income ranks are not, interacting exposures with p_i in small populations (such as children with siblings in rare location sequences) makes residualization on sequence-specific slopes $\hat{\boldsymbol{\beta}}^1$, and the resulting estimates $\hat{\boldsymbol{\mu}}^0$ and $\hat{\boldsymbol{\mu}}^1$, very unstable.

it is not a problem that identification of slopes $\hat{\boldsymbol{\beta}}^1$, and $\hat{\boldsymbol{\mu}}^1$ comes only from variation in exposures at different parent ranks. Residualization on $\boldsymbol{\beta}^0 \cdot \boldsymbol{h}_i + \boldsymbol{\beta}^1 \bar{p}_{f_i \boldsymbol{d}_i} \cdot \boldsymbol{s}_i = (\boldsymbol{\kappa}^0 + \boldsymbol{\kappa}^1 \bar{p}_{f_i \boldsymbol{d}_i}) \cdot \boldsymbol{s}_i + \bar{\theta}_{\boldsymbol{h}_i \bar{p}_{f_i \boldsymbol{d}_i} \boldsymbol{s}_i}$ incorporates both causal effects $\boldsymbol{\kappa}$ and selection effects $\bar{\theta}_{\boldsymbol{h}_i \bar{p}_{f_i \boldsymbol{d}_i} \boldsymbol{s}_i} = \mathrm{E}[\theta_i | \boldsymbol{h}_i, \bar{p}_{f_i \boldsymbol{d}_i} \boldsymbol{s}_i]$, and $\tilde{\epsilon}_i$ captures within-family residual variation. This estimator absorbs sequence and family effects, identifies exposure effects from within-family variation in the ages at which children move, and captures heterogeneity related to family-average parent ranks.

For comparison with Estimator A1, we maintain the weaker identification assumption

Assumption A2. Conditional on family-by-sequence intercepts, $\boldsymbol{h}_i$, and sequence-by-parent-rank slopes, $\bar{p}_{f_i \boldsymbol{d}_i} \boldsymbol{s}_i$, exposures and their interactions with parent family income rank $\boldsymbol{e}_i$ and $\bar{p}_{f_i \boldsymbol{d}_i} \boldsymbol{e}_i$ are jointly orthogonal to residual variation in childrens' outcomes

$$\mathrm{E}[\tilde{\epsilon}_i | \boldsymbol{e}_i, \bar{p}_{f_i \boldsymbol{d}_i} \boldsymbol{e}_i, \boldsymbol{h}_i, \bar{p}_{f_i \boldsymbol{d}_i} \boldsymbol{s}_i] = \mathrm{E}[\tilde{\epsilon}_i | \boldsymbol{h}_i, \bar{p}_{f_i \boldsymbol{d}_i} \boldsymbol{s}_i]$$

Assumption A2 only requires that there should be no selection within families and location sequences. For example, conditional on moving to Los Angeles from New York, older siblings should not have better unobservables than younger siblings.

Comparison of Estimators A1 and A2. We test for bias in Estimator A1 under the maintained assumption that Estimator A2 is unbiased. Formally, we define bias as the difference $\hat{\boldsymbol{\mu}}^{A1}_{\bar{p}} - \hat{\boldsymbol{\mu}}^{A2}_{\bar{p}}$, and test for significant bias for all CZs jointly and for each CZ separately.

We implement coefficient comparison tests as recommended by Pei, Pischke, and Schwandt (2017), estimating both models on a common population and accounting for shared observations, by seemingly-unrelated estimation.[23]

First, we test the joint null hypothesis that all estimates from Estimator A1 are unbiased, with $\hat{\boldsymbol{\mu}}^{A1}_{\bar{p}} - \hat{\boldsymbol{\mu}}^{A2}_{\bar{p}} = \mathbf{0}$, by the Wald statistic

$$H_{\bar{p}} = (\hat{\boldsymbol{\mu}}^{A1}_{\bar{p}} - \hat{\boldsymbol{\mu}}^{A2}_{\bar{p}})' (\hat{\boldsymbol{V}}_{\hat{\boldsymbol{\mu}}^{A1}_{\bar{p}} - \hat{\boldsymbol{\mu}}^{A2}_{\bar{p}}})^{-1} (\hat{\boldsymbol{\mu}}^{A1}_{\bar{p}} - \hat{\boldsymbol{\mu}}^{A2}_{\bar{p}})$$

and find $H_{25} = XXXX$ and $H_{75} = XXXX$ which, against critical value $\chi^2_{0.05,761} = 826$, suggest bias in at least one location.

Second, to determine where bias might be relevant, we test the null hypothesis for each CZ that its estimate from estimator A1 is unbiased, with $\hat{\mu}^{A1}_{c\bar{p}} - \hat{\mu}^{A2}_{c\bar{p}} = 0$, by the t-statistic. Figure **??** shows that the bias in Estimator A1 is significant only in a small number of

[23] We stack two copies of the dataset, run one regression on each, and obtain the combined VCE matrix from the Stata command `suest`, clustering on paired observations. Following the design-based interpretation of Abadie et al. (2020), we do not correct for degrees of freedom lost to sequence fixed effects, but we do correct for degrees of freedom lost to family fixed effects within sequences.

locations: nearly all differences have a t-statistic below 1.960, and accounting for multiple testing with a Bonferroni correction, all but X have a t-statistic below 3.993. Two of the X are Los Angeles and New York, suggesting caution in comparisons with these CZs. These exceptions reinforce the value of testing Assumption A1, with a focus on location-specific selection.

Overall, the evidence suggests that Assumption A1 is violated in a handful of locations, without implicating other locations.

4.3.2 Child-Level Factors

We are currently testing for potential child-level confounds.

5 Additional Results and Discussion

5.1 Observed Outcomes

In this section, we compare the exposure effect estimates with children's observed outcomes. This comparison is valuable for two reasons. First, observed outcomes are currently the preferred choice to make policy recommendations (e.g. Bergman et al., 2024; Chetty et al., 2026) because they are less noisy than exposure effects, but it is not known whether bias in observed outcomes makes them worse predictors of policy-relevant causal effects. Second, it is not known whether selection into locations (specifically, location sequences, not location exposures) reinforces or compensates for exposure effects.

5.1.1 Selection into Location on Family- and Child-Specific Factors

We find that there is substantial bias in observed outcomes due to selection on unobserved family- and child-specific factors. These results are shown in Figure **??** and **??**, which map the selection term in observed outcomes by CZ, at the 25th and 75th percentiles of parent family income. Family- and child-specific factors are controlled in exposure effect estimates, by location-sequence fixed effects. However, they are not controlled in specifications that omit location-sequence fixed effects, both for the population of movers and for the population of stayers. The bias in observed outcomes makes many large cities, such as New York, Los Angeles, and San Francisco, appear much better for children than they are in a causal sense; and it makes much of the South and Southwest appear much worse for children than they are in a causal sense. These findings suggest that policy recommendations based on observed outcomes should be avoided where sufficiently-precise exposure effect estimates are available.

5.1.2 Assortative Migration

We find no evidence for selection bias in observed outcomes that reinforces or compensates exposure effects. We measure selection bias by estimating exposure effects without and with fixed effects for the sequence of moves, where estimates without fixed effects can come from either the movers population or the stayers population, and then subtracting the estimates with fixed effects from the estimates without fixed effects. These results are shown in Figure **??** and **??**, which scatter the selection term against the exposure effects, at the 25th and 75th percentiles of parent family income, and fit a regression line weighted by exposure shares. While regressing this estimated difference on the estimates with fixed effects should theoretically result in a negative coefficient estimate due to negatively correlated errors, the estimates are precise enough that this does not affect the regression in practice, and we do not find a consistently negative slope.

The observation that parents do not selectively move toward places with positive exposure effects is important for two reasons. First, if parents are not choosing where to move based on exposure effects and their own unobservables, then they are probably also not choosing when to move, buttressing Assumption A1. Second, if parents are not choosing where to move based on exposure effects in general, then there must be some reason why, setting up our discussion of revealed preference in Section 5.5.

5.2 Forecasts

We calculate MSE-minimizing forecasts by shrinking estimates toward observed outcomes (and towards zero, where both are noisy). We find that only the most noisy estimates are meaningfully shrunk toward observed outcomes. Under the assumption that exposure effects are normally distributed, or with very high signal-to-noise ratios such as we find in Section 4.2, Armstrong, Kolesar, and Plagborg-Moller (2022) suggest that the RMSE can be used to form confidence intervals as usual. Figure **??** and **??** display the forecasts and confidence intervals based on their root-mean squared errors. Given that the shrunk estimates are much less variable than the raw exposure effect estimates, we map the shrunk estimates in Figure **??** and **??**; we find that geographic patterns remain stable.

We compare the shrunk estimates with their weighted components, the raw exposure effect estimates and the observed outcomes, in Figure **??**. Overall, the forecasts are very similar to estimated exposure effects, and very different from observed outcomes. Figure **??** and **??** map the discrepancies between forecasts and the underlying data, showing where observed outcomes over- and under-estimate the effects of spending more time in a particular CZ in childhood: these results emphasize that the forecasts closely match the exposure effects,

and deviate from the observed outcomes according to the pattern of selection discussed in Section 5.1. Our results suggest that observed outcomes, and forecasts based on less-precise exposure effect estimates, are not reliable and should not be used to make recommendations.

5.3 Pairwise Comparisons

Because decisions about when and where to move depend on exposure effects relative to other CZs, we measure differences $\hat{\mu}_{cp} - \hat{\mu}_{c'p}$, and their uncertainty, by combining estimates from equation (5) and (6),

$$\hat{V}_{\hat{\mu}_{cp}-\hat{\mu}_{c'p}} = \hat{V}_{\hat{\mu}_{cp}} + \hat{V}_{\hat{\mu}_{c'p}} - 2\hat{V}_{\hat{\mu}_{cp},\hat{\mu}_{c'p}} \tag{11}$$

For example, for children moving from New York to Los Angeles, with parents at the 25th percentile of parent family income, spending one more year in New York instead of Los Angeles would result in child family income $\hat{\mu}_{NY25} - \hat{\mu}_{LA25} = -X.XXXX$ $(X.XXXX)$ percentiles lower. At the 75th percentile of parent family income, the same year in New York would result in child family income $\hat{\mu}_{NY75} - \hat{\mu}_{LA75} = -X.XXXX$ $(X.XXXX)$ percentiles lower. These effects are typical in size: at the 25th and 75th percentiles of parent family income, XX and XX percent of pairwise effects, respectively, are larger than 0.0417 percentiles per year of exposure; scaled to full exposure, these effects amount to a 1 percentile change in children's family income due to choosing their origin CZ over their destination CZ, or vice versa.

How informative are pairwise comparisons of our estimates, for the typical family deciding when and where to move? What testing criteria should we use? Family decisions about when to move involve single comparisons. Confidence intervals around the benefit of the move, in terms of child percentile ranks gained per year of exposure, can be constructed based on a standard critical value of 1.96. These single comparisons between CZs are the most likely use of the estimates, and the use most consistent with the estimation strategy, which is based on variation in the timing of moves between CZs rather than variation in where families move.

Figure **??** presents a histogram of absolute t-statistics across all pairwise comparisons, weighted by the number of children making each pairwise move with their families. We find that at the 25th and 75th percentiles of parent family income, XX and XX percent of moves, respectively, would produce a significant gain or loss in children's outcomes, at conventional significance levels relevant for deciding when to move. Informally, using a 1 percentile cutoff for economic significance, XX percent of economically significant differences are statistically significant at conventional levels. Conversely, among economically insignificant differences, only XX percent are statistically significant at conventional levels.

Family decisions about where to move involve multiple comparisons, of the family's origin CZ with up to 760 potential destination CZs, and are likely to involve identifying the best potential destination among those considered (a ranking problem as in Mogstad et al., 2024; discussed in Section 5.4 below) or identifying the gain from moving to the best potential destination (a selection problem as in Andrews, Kitagawa, and McCloskey, 2024). To give a rough idea of how many pairwise differences are significant with multiple comparisons, we consider a conservative Bonferroni correction: the critical value controlling the per-family error rate (PFER) is 3.99. By this criterion, we find that XX percent of moves would produce a significant gain or loss. Of course, because many families consider only a few potential destinations, the appropriate critical value may be much smaller.

Are comparisons informative for all CZs, or only some CZs with extreme estimates? Figure **??** and **??** report the fraction of tests that exceed the conventional critical value of 1.96, and the corrected critical value of 3.99, for each CZ. At conventional levels we find that two-fifths or more of comparisons for a given CZ are significant, and more than four-fifths for most of the largest urban CZs. Only rarely does the fraction of significant comparisons fall below one-fifth. That is, from most CZs, most pairwise comparisons are informative for families considering when to move. Turning to tests for comparisons across all potential destinations, most of the most-populous CZs reach at least one-fifth to two-fifths of comparisons significant, with others reaching above four-fifths significant; however, it becomes more common to find origin CZs where little can be said about what pairwise differences are significant. We view this latter result as a lower bound on informativeness.

We conclude that most families making decisions based on the estimates, regardless of where they live, can consider the estimates informative both in single comparisons and multiple comparisons. However, it remains to be seen how the estimates perform when used as inputs to compound problems including ranking and selection.

5.4 Rankings

Mogstad et al. (2024) report that the ranking in Chetty and Hendren (2018b) is uninformative due to estimation noise and multiple testing. Although our estimates address the problem of estimation noise, it is not a given that hypothesis tests ranking CZs based on our estimates will meet the stringent critical values set by Mogstad et al.'s controls for family-wide error rates (FWER). Thus, we apply the ranking method in Mogstad et al. (2024) to our estimates; because they did not have access to full VCE estimates from Chetty and Hendren (2018b), we modify their replication code to compute standard errors for pairwise comparisons using the covariance estimates, finding a small sharpening of the ranking.

Ranking results are presented in Figure ?? and ??, including "marginal" confidence sets for each CZ relative to all the other CZs, and "simultaneous" confidence sets for all possible pairwise comparisons. Overall, our new estimates are precise enough to deliver a highly-informative ranking.

We observe that FWER control for marginal comparisons between a given origin CZ and 50 or more destination CZs is conservative relative to how a typical user would answer questions about where to move based on exposure effects. It is harder to think of a thought experiment where FWER control for simultaneous comparisons between all possible origin and destination CZs is appropriate, but a possible example is where a decisionmaker is trying to decide which origin CZ's residents to target for a policy recommending moves to one or a few of the possible destination CZs. This underscores the fact that our estimates will be even more reliable for the more-limited comparisons that most users would be likely to make.

5.5 Revealed Preference

We use flows of families between locations to identify parents' revealed preferences for places. Formally, we rank locations by network centrality, as in Google's PageRank (Page, 1996) algorithm: each move from one place to another is a weighted (by the rank of the origin) vote for the destination. Figure ?? shows that parents have strong revealed preferences for places with negative exposure effects, setting up an important tradeoff between parents' and children's best interests. We also find evidence of strong positive selection on families' unobservables into the places that parents prefer: Figure ?? plots the selection term in observed outcomes against parents' revealed preference. Finally, we map parents' revealed preference in Figure ??. The places that parents prefer are generally major urban centers, for example New York, Newark, Chicago, San Jose, San Diego, Los Angeles, and New Orleans.

5.6 Mechanisms

Locations with positive exposure effects are also locations with negative parent earnings premiums, and lower rents and house prices. Figure ?? and ?? plot median rents and earnings premiums by CZ against exposure effects, finding strong negative gradients for both. These results are consistent with the fact that parents don't prefer these locations. Figure ?? and ?? plot median rents and earnings premiums by CZ against parents' revealed preferences, finding positive relationships for both: parents prefer places with high rents (which capture the value of those places to renters) and high earnings premia (which justify the rents).

Thus, opportunity for parents comes "dressed in overalls and looks like work," but opportunity for children is a bargain only when it comes to rent. This tradeoff presents a serious challenge for policies that would move whole families "to opportunity."

5.7 Heterogeneity

We impute parents' nativity to estimate exposure effects for the native children of immigrant parents. We find informative exposure effects, which are however somewhat noisier than the full-population estimates. Figure **??** presents the estimates for immigrants in the top 50 CZs by population. These estimates make it clear that more precision would help with interpretation, especially for smaller CZs. Carrying out shrinkage as in Section 5.2, and adding covariates to improve precision, we find relatively precise estimates that still load primarily on the causal estimates, as shown in Figure **??**. Figure **??** and **??** present the shrinkage estimates for the top 50 CZs, and for all CZs, showing the improvement in precision.

With greater precision, we are able to map exposure effects for the children of immigrants, showing in Figure **??** and **??** that coarse geographic patterns of mobility are similar to those for the children of natives. Estimates are more precise for the 25th percentile of parent family income, because the immigrant population below median income is roughly three times larger than the immigrant population above median income. Many estimates are significant at conventional levels, and hypothesis tests for larger CZs often exceed Bonferroni-corrected critical values, confirming these geographic patterns.

In Figure **??** and **??**, we use our estimates to rank the top 25 CZs by immigrant population, and find similar patterns of exposure effects to the full population.

Like all parents, immigrant parents have revealed preferences for places that have negative exposure effects for their children. Figure **??** maps parents' revealed preferences by CZ. We will confirm the role of earnings premia for revealed preferences.

6 References

Abramitzky, Ran, Leah Boustan, Elisa Jacome, and Santiago Perez. 2021. "Intergenerational Mobility of Immigrants in the United States over Two Centuries," *American Economic Review* 111:2, pp. 580-608

Alexander, J. Trent, David Bleckley, Jonathan Fisher, Katie Genadek, Susan Hautaniemi Leonard, and Aristotle Magganas. 2024. "Newly-Available Individual-Level U.S. Tax Data from 1969-1994," ADEP Working Paper 2024-02

Aloni, Tslil and Hadar Avivi. 2025. "One Land, Many Promises: Assessing the Consequences of Unequal Childhood Location Effects," working paper

Amornsiripanitch, Natee, Paul Gompers, George Hu, and Kaushik Vasudevan. 2023. "Getting schooled: Universities and VC-backed immigrant entrepreneurs," *Research Policy* 52:7

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

Armstrong, Timothy, Michal Kolesár, and Mikkel Plagborg-Møller. 2022. "Robust Empirical Bayes Confidence Intervals," working paper

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

Bernstein, Shai, Rebecca Diamond, Abhisit Jiranaphawiboon, Timothy McQuade, and Beatriz Pousada. 2025. "The Contribution of High-Skilled Immigrants to Innovation in the United States," NBER Working Paper No. 30797

Card, David, Jesse Rothstein, and Moises Yi. 2025. "Location, Location, Location," *American Economic Journal: Applied Economics* 17:1, pp. 297-336

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," *Quarterly Journal of Economics* 129:4, pp. 1553–1623

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

Chetty, Raj and Nathaniel Hendren. 2018b. "The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates," *Quarterly Journal of Economics* 133:3, pp. 1163-1228

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

Chetty, Raj, John Friedman, Nathaniel Hendren, Maggie Jones, and Sonya Porter. 2026. "The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility," *American Economic Review* 116:1, pp. 1-51

Cookson, J. Anthony, Benedict Guttman-Kenney, and William Mullins. 2025. "Immigration and Credit in America," working paper

Kim, Jae Kwang and Wayne Fuller. 2004. "Fractional hot deck imputation," *Biometrika* 91:3, pp. 559-578

Kim, Jae Kwang and J.N.K. Rao. 2009. "A unified approach to linearization variance estimation from survey data after imputation for item nonresponse," *Biometrika* 96:4, pp. 917-932

Klopfer, John and Corbin Miller. 2024. "Using SSNs to Identify Immigration, Selection, and Censoring in Administrative Data," HKU Jockey Club Enterprise Sustainability Global Research Institute Paper No. 2024/009

Miller, Douglas, Na'ama Shenhav, and Michel Grosz. 2023. "Selection into Identification in Fixed Effects Models, with Application to Head Start," *Journal of Human Resources* 58:5, pp. 1523-1566

Mogstad, Magne, Joseph Romano, Azeem 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, pp. 476-518

Pei, Zhuan, Jörn-Steffen Pischke, and Hannes Schwandt. 2017. "Poorly Measured Confounders are More Useful on the Left Than on the Right," NBER Working Paper No. 23232

Ramakrishnan, S. Karthik. 2004. "Second-Generation Immigrants? The '2.5 Generation' in the United States," *Social Science Quarterly* 85:2, pp. 380-399

Sorkin, Isaac. 2018. "Ranking Firms Using Revealed Preference," *Quarterly Journal of Economics* 133:3, pp. 1331-1393

White, Halbert. 1996. *Estimation, Inference, and Specification Analysis*. Cambridge: Cambridge University Press

Yonker, Scott. 2017. "Geography and the Market for CEOs," *Management Science* 63:3, pp. 587-900