One Land, Many Promises: Assessing the Consequences of Unequal Childhood Location Effects

Tslil Aloni   Hadar Avivi*

July 21, 2024

Newest version here

Abstract

This paper estimates the causal effects of childhood residential location on the adult income of native-born Israeli children and the children of immigrants from the former Soviet Union and studies the consequences of location effect heterogeneity on the design and effectiveness of neighborhood recommendation policies. The causal effects of childhood location contribute substantial variability to the adult earnings of Israeli children. While the place effects of high-income immigrants and high-income natives are strongly correlated, location effects for low-income immigrants are uncorrelated with location effects for low-income natives. Guided by these findings, we develop a policy targeting framework aiming to recommend the top locations in Israel while incorporating the constraint that the policymaker cannot make ethnicity-dependent location recommendations. Using empirical Bayes tools, we find that targeting policies based on pooled population-wide averages yield inferior outcomes for immigrants. Robust targeting strategies designed to perform well against the least favorable sorting patterns reveal a set of 10 cities that are likely to benefit children of both groups.

*Hadar Avivi: Economics Department, U.C. Berkeley. Email: havivi@berkeley.edu. Tslil Aloni: Economics Department, New York University. Email: aloni@nyu.edu. We are grateful to Patrick Kline, Christopher Walters, and Hilary Hoynes for their invaluable guidance and support on this project. For very helpful comments, we would like to thank Livia Alfonsi, Sydnee Caldwell, Luisa Cefala, Nick Flamang, Conrad Miller, Enrico Moretti, Ben Scuderi, Yotam Shem-Tov, and Damián Vergara. We thank the seminar participants at UC Berkeley, the University of Chicago Economics, the University of Chicago Harris School of Public Policy, Yale, the London School of Economics, University College London, Tel-Aviv University, the Hebrew University, Ben-Gurion University, and Richman University. We are grateful to Avigail Sageev for outstanding research assistance and to Anat Katz, Taina Trachtenberg, and Rivka Azulay from the Israeli Central Bureau of Statistics for invaluable administrative support. We are grateful to the Chief Economist at the Israeli Ministry of Finance and to Assaf Geva for supporting the research financially and conceptually. Hadar gratefully acknowledges financial support from the Institute for Research on Labor and Employment and the Opportunity Lab at UC Berkeley.

1 Introduction

A growing body of literature finds that childhood locations have a significant and long-lasting effect on outcomes in adulthood (see Chyn and Katz, 2021, for a review). This evidence is the basis for "moving to opportunity" policies aiming to encourage low-income housing voucher recipients to move to high-opportunity neighborhoods (Katz et al., 2001; Bergman et al., 2019). This research has focused entirely on unified policies that provide a single recommendation to a wide group of families using neighborhood-level estimates. The effectiveness of such unified policies relies on either limited treatment effect heterogeneity or prior knowledge about recipients' behavioral responses. While the literature has emphasized the optimality of personalized treatment rules (Kleinberg et al., 2018b; Cowgill and Tucker, 2019; Rambachan et al., 2020), the research on the possibilities of restricted unified policies is limited.

In this paper, we engage with this question by exerting two main efforts. First, we provide striking evidence that childhood location effects vary substantially for low-income children with different backgrounds. Using a comprehensive administrative dataset from Israel, we establish that, similar to Chetty and Hendren (2018a,b), place of birth contributes substantial variability to the adult earnings of both native-born and immigrant children. However, the correlation between these effects on low-income immigrant and native-born children is close to zero. Following these findings, we then study the potential consequences of this heterogeneity on the outcomes of potential recipients of neighborhood recommendation policies and propose an alternative unified recommendation policy that accounts for heterogeneity.

Our paper begins by revisiting the benchmark estimates of childhood location effects from Chetty and Hendren (2018a,b) in Israel and estimates effects separately for immigrants from the former Soviet Union and native-born children. Focusing on children born between 1980 and 1991, we estimate the effects on income rank at age 28 for each city and regional council.[1] By exploiting the mass migration wave from the former Soviet Union to Israel between 1989 and 2000, during which 1 million immigrants arrived and spread throughout the country, we can identify location effects in most major cities for both groups.

[1] A regional council is a group of small localities, such as small towns or *Kibutzim*, that are geographically close and share the same local governing council. Cities and regional councils are the smallest local government units in Israel.

We identify causal location effects by leveraging variations in children's exposure time to different cities during childhood due to household moves at different ages. This strategy combines variations in the timing of moves across locations within Israel and the age at which children migrated from the former Soviet Union. This strategy does not require families to sort randomly but assumes that among families with the same location choices, the child's age at arrival is unrelated to unobserved components affecting potential outcomes. To support this, we demonstrate that these location choice fixed effects balance observable family characteristics across children's ages upon arrival to each city, suggesting unconfounded comparisons. Additionally, our model assumes that location effects are linear with exposure time. We justify this by demonstrating that the standard diagnostics in the literature indicating a linear relationship between exposure time and mean outcomes of children who spent all their childhood in the same location also hold in Israel (Chetty and Hendren, 2018a; Deutscher, 2020; Laliberté, 2021) and provide additional evidence that the location effects themselves are linear with exposure time.

Childhood location effects vary substantially for both native-born and immigrant children. To quantify the extent of across-city heterogeneity, we estimate the standard deviation of location effects for natives and immigrants, adjusting for sampling error. For a child with parents at the 25th percentile of the national income distribution, a one standard deviation increase in city quality for a single year boosts income at age 28 by 0.44% for natives and 0.5% for immigrants per year, compared to the mean. Extrapolating over 18 years of childhood, growing up in a one standard deviation better city from birth would increase the income in adulthood of a native-born child by 8% and of an immigrant child by 9%.

Childhood location effects also vary substantially within cities across immigration groups, with a pattern that differs by household income. We find that the correlation between the location effects of immigrants and natives among low-income families at the 25th percentile of the income distribution is close to zero, while there is a strong positive correlation between the location effects of immigrants and natives among high-income families. This result implies that there is no single "promised land" for low-income families, i.e., places that generate high adult income for one group do not generally boost income for the other. We show that this zero correlation is not driven by differences in high school attendance patterns within locations, within-city heterogeneity in neighborhood effects, or mismeasurement of immigrant parental income.

Large, diverse cities with a substantial share of both immigrants and natives are more likely to benefit immigrants. This finding could potentially reconcile the mixed results in the literature on the effects of the geographic concentration of immigrants and refugees on their outcomes. While Edin et al. (2003) and Beaman (2012) find that immigrant enclave size positively affects immigrants' outcomes in settings with a low city-level immigrant share, Abramitzky et al. (2020) points out that leaving large Jewish communities in New York in the 20th century was beneficial for Jewish immigrants. Moreover, places with high crime rates and high municipal welfare expenditure per capita are more likely to be detrimental to native-born children, while these measures are less predictive of low-income immigrant location effects. Previous literature has emphasized the relationship between poverty-related covariates and location effects, using these characteristics to target housing policy (Katz et al., 2001). Our findings suggest that such targeting strategies may not be useful for immigrants in the Israeli context.

Motivated by these findings, we next study the consequences of heterogeneity on the policy implemented in the Creating Moving to Opportunity (CMTO) experiment (Bergman et al., 2019), which provided housing voucher recipients with recommendations on where to move based on tract-level upward mobility estimates. We focus on a unified policy that provides the same recommendations to all groups. Although the literature suggests that the optimal policy should ideally be personalized and based on group identity (Chan and Eyster, 2003; Cowgill and Tucker, 2019; Rambachan et al., 2020; Ellison and Pathak, 2021), this restriction is motivated by legal and moral constraints in many countries, where it is unacceptable to base public programs on ethnic identity or promote segregation.

Using a decision-theoretic framework, we work with a Bayesian decisionmaker who acknowledges uncertainty about the true location effects and makes decisions based on expected rather than true loss. Our analysis employs an empirical Bayes approach, where forecasts and decisions are based on the posterior distribution of location effects, treating the estimated distribution as a prior. Our model involves a decisionmaker who evaluates each decision rule relative to the first-best personalized policy, aiming to minimize the *regret* of not using it (Savage, 1954; Manski, 2004). This approach, motivated by discussions with Israeli policymakers, addresses the practical issue in our empirical model where each group's location effects are identified only up to a base-level normalization.

A policy that ranks locations based on a pooled average estimate of city quality (similar to that considered by Bergman et al. (2019)) results in lower weights on the gains for minority groups, producing inferior outcomes for such groups. These unequal outcomes arise from two sources. First, the decisionmaker's inability to target treatment by ethnic group ex-ante prevents the policy from leveraging the heterogeneity in location effects across groups. Second, the decision maker's ambiguity regarding which households will respond to each particular policy recommendation and how. With treatment effect heterogeneity, some compliance behavior with the policy may dilute its effectiveness if the gains for households that respond are very different from the overall average effect.

We suggest an alternative targeting policy that accounts for these issues by providing a list of recommended locations optimal under the least favorable compliance scenario. This is a *minimax* strategy (Wald, 1950) which minimizes the maximal loss under the worst possible behavioral response. We show that this robust policy can generate substantial advantages for minority groups and achieve more equitable outcomes. With the minimax policy, we can pinpoint at least ten cities that offer benefits for both groups, where the worst-case outcome for either group is 40% better than under the city-level average policy. Additionally, we can ensure that, on average, no more than 10% of the recommended cities would yield outcomes inferior to those resulting from the current status quo sorting patterns.

This paper contributes to several threads of literature. First, we add a new perspective to the vibrant discussion on the challenges that might arise from the neighborhood recommendation policies proposed in the CMTO experiment. So far, the literature has focused primarily on issues of identification (Heckman and Landersø, 2021; Eshaghnia, 2023), measurement (Chen, 2023; Aliprantis et al., 2024), and inference (Andrews et al., 2022; Mogstad and Torsvik, 2021; Mogstad et al., 2023), where the latter work emphasizes the ramifications of ranking locations based on noisy estimates rather than their true values. Although Chetty et al. (2018) acknowledge the potentially multifaceted nature of locations, the literature has not considered the complications it generates. In the CMTO, for example, there is no guarantee that all recommended places are indeed beneficial to all participants. While Mogstad et al. (2023) mention this concern regarding the risk of forming policy based on noisy estimates, similar logic applies also if the signal varies. In this paper, we directly address the policy implications of location effect heterogeneity by modeling the uncertainty from both heterogeneity and unknown compliance, along with the uncertainty driven by measurement error.

Methodologically, our work relates to a growing literature on Empirical Bayes ranking and prediction methods that use shrinkage estimates to identify the value added of schools, teachers, hospitals, and discriminatory firms (Chetty et al., 2014a; Abdulkadiroğlu et al., 2020; Abaluck et al., 2021; Kline et al., 2022). Recent work in econometrics has emphasized that such tasks are analogous to multiple testing problems, where decisions result from constraints on various sorts of error rates (Gu and Koenker, 2020; Kline and Walters, 2021; Kline et al., 2023; Mogstad et al., 2023). We add to this literature by modeling the risk a decision maker faces, distinguishing between the risk stemming from effect heterogeneity, unknown behavioral responses, and statistical noise. As such, our paper contributes to the literature on optimal statistical treatment rules (Manski, 2004; Kitagawa and Tetenov, 2018; Manski, 2021). Similar to Christensen et al. (2022), our model departs from classic approaches to these problems by considering how optimal decisions depend on partially identified parameters. In our setting, location effects are point-identified (so the decision maker faces only statistical uncertainty), while household compliance patterns are not, creating ambiguity regarding the final allocation of true payoffs across families.

Additionally, this paper extends a growing literature in economics on algorithmic bias and fairness (Kleinberg et al., 2018b; Cowgill and Tucker, 2019; Rambachan et al., 2020; Liang et al., 2021), and the equity-efficiency tradeoffs of affirmative action programs (Lundberg, 1991; Chan and Eyster, 2003; Ellison and Pathak, 2021), where papers in both strands conclude that the optimal policy should exploit all available information, including group identity variables. Instead, we explore the possibilities for a policy conditional on a suboptimal restricted algorithm, which, to our knowledge, hasn't been studied before. Our model demonstrates that we can improve the fairness of the restricted policy by modeling the uncertainty generated by such restrictions using a decision-theoretic framework. This approach can be extended to settings with anti-discriminatory laws or group-directed treatments, such as teacher and school assignments (Abdulkadiroğlu et al., 2020; Biasi et al., 2021; Bobba et al., 2021; Rose et al., 2022; Bates et al., 2022; Graham et al., 2023), admission policies (Ellison and Pathak, 2021), job training programs (Card et al., 2018), criminal justice

(Chouldechova, 2017; Kleinberg et al., 2018a; Agan and Starr, 2018; Ba et al., 2021), and recruiting decisions (Li et al., 2020).

Lastly, we contribute to the literature on the causal effects of locations on immigrants. Previous studies find that immigrant adults benefit from clustering geographically (Bartel, 1989; Damm, 2009; Bertrand et al., 2000; Munshi, 2003; Edin et al., 2003; Card, 2009; Buchinsky et al., 2014; Dustmann et al., 2016; Altonji and Card, 2018; Munshi, 2020). We contribute by providing additional evidence that locations matter for immigrant children (Gould et al., 2004, 2011; Damm and Dustmann, 2014; Abramitzky et al., 2020) and by quantifying the city-level immigrant-native causal location effect gap.

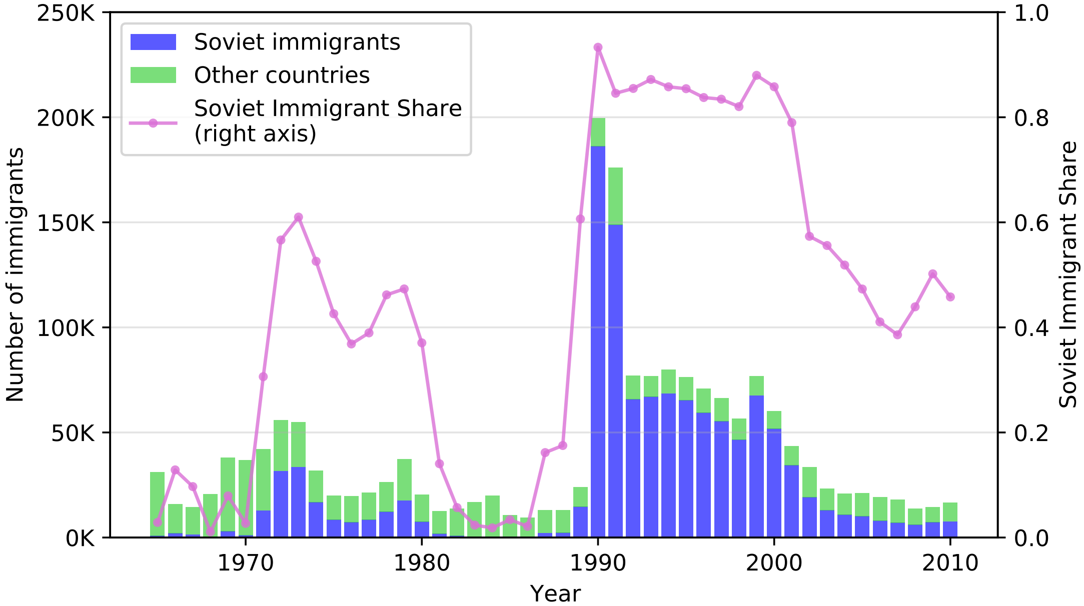
The remainder of this paper is organized as follows: Section 2 covers the historical context. Section 3 outlines the empirical model. Section 4 describes our data. In Section 5, we estimate the extent of variation in causal location effects in Israel, both across cities and between immigrants and natives. Predictors of location effects are explored in Section 6. Section 7 provides evidence of the robustness of our approach using alternative specifications. In Section 8, we estimate parametrically the joint distribution of location effects. Section 9 analyzes policy design in the presence of place effect heterogeneity, and Section 10 concludes.

2 Historical Context

In 1989, the former Soviet Union relaxed its emigration restrictions, resulting in one of the most significant human movements of the late 20th century. Prior to this relaxation, restrictive emigration laws and tight governmental controls made it nearly impossible for Soviet residents to leave the country. As the USSR disintegrated, these legal barriers dissolved, and approximately 7 million Soviet residents left the Soviet Union between 1989 and 2000 (Abramitzky et al., 2022). Among them, more than 1 million Jewish immigrants arrived in Israel, increasing Israel's population by 20%.

Figure 1 presents the number of Soviet immigrants entering Israel by year. The bulk of the migration wave, over 300 thousand immigrants, arrived in a relatively short time span, between 1989 and 1991, which accounted for 7 percent of the Israeli population prior to the immigration. The peak in the first years was followed by a consistent flow of 60 thousand immigrants a year, which lasted the entire decade. In total, more than one million Jews left the former Soviet Union, which amounted to around a fifth of the Israeli population of 1989.

Figure 1: Annual number of Former Soviet Union immigrants and other countries to Israel



Note: This figure displays the number of migrants to Israel between 1965 and 2019 arriving from the Soviet Union and other countries. On the right axis, the pink line displays the fraction of Soviet immigrants. Source: the Israeli Central Bureau of Statistics.

Upon their arrival, Soviet Jews were granted full citizenship. Thus, like any other citizen, they had unrestricted access to social services, education, health care, and social security benefits. The Soviet immigrants were free to choose where they wanted to live in Israel and faced no formal labor market restrictions (Buchinsky et al., 2014). The Israeli government supported immigrants in various forms, including a modest grant (called the "absorption basket") lasting for one year after arrival, free Hebrew classes, and local centers that provided initial information and support to facilitate integration.

This migration wave comprises several favorable features for studying the role of parental location choices on children's long-run economic outcomes. First, it is a massive and non-restrictive migration wave, with entire families immigrating together. This unique feature allows us to identify causal location effects separately for Soviet immigrants in multiple locations. Second, the lifting of Soviet emigration restrictions was unexpected. Therefore, the timing of the migration wave is likely exogenous, and therefore, children's age of migration is plausibly unrelated to their parents' preferences and characteristics. Lastly, immigrants were granted Israeli citizenship upon arrival, thus facing no differential regulatory restrictions compared to natives. Therefore, any immigrant-native gap we find cannot be explained by institutional restriction or regulation.

3 Empirical Model

3.1 Conceptual Framework

Consider a population of children indexed by i and a set of locations indexed by $j \in \{1, ..., J\}$. Let $Y_i(e)$ denote child i's potential adult income as a function of the number of years of exposure to each location, represented by the vector $e = (e_1, ..., e_J)'$. We assume that childhood locations affect children's long-run outcomes from birth to age 18, with e_j representing the number of years of exposure to city j before age 18 such that $\sum_j e_j = 18$.[2] We model these potential outcomes with an additive structure:

$$Y_i(e) = \sum_{j=1}^{J} \theta_{j\chi(i)} \cdot e_j + \xi_i(e), \tag{1}$$

where $\theta_{j\chi}$ represents the contribution to adult income of an extra year in city j to child i whose observed characteristics are $\chi(i)$, and $\xi_i(e)$ is the error term. This specification rules out location effect heterogeneity by child's age or complementarity or substitutability between time spent in different places. The error term, $\xi_i(e)$, represents all the other age-, time-, or location-dependent shocks beyond the variation by exposure time and childhood city that affect children's long-run outcomes, such as time-invariant and time-varying parental investments, moving costs, or age-specific shocks. The observed outcome for child i is given by $Y_i = Y_i(E_i) = \sum_j \theta_{j\chi(i)} E_{ij} + \xi_i(E)$, where $E_i = (E_{1i}, ..., E_{Ji})'$ represents child i's realized years of exposure to each city from birth to age 18.

Each child i belongs to a group $g(i) \in \{\mathcal{N}, \mathcal{I}\}$, either natives ($\mathcal{N}$) or immigrants ($\mathcal{I}$), and is characterized by a parental income rank variable $p(i)$. Given these characteristics, we

[2] In Appendix Section D, we present evidence suggesting that in Israel, the last year in which cities affect economic outcomes is approximately 18. This finding aligns with the Israeli institutions in which most children enlist in the army immediately after high school.

parameterize the annual effects of childhood location as follows:

$$\theta_{j\chi(i)} = \underbrace{\alpha_{jg(i)} + \eta_{jg(i)} \cdot p(i)}_{\equiv \theta_{jg(i)p(i)}}. \tag{2}$$

This linear specification follows earlier work indicating that a linear relationship between parent income rank and location effects provides a good empirical approximation (Chetty et al., 2014b).[3] The intercept α_{jg} measures the effect of spending one more year in city j for a child of group g whose parental income is at the lowest percentile in the national income distribution, and the slope η_{jg} measures the one-year return to parental income in location j for a child who belongs to group g. Therefore, the total one-year location effect in city j for a child of group g with parents' income p is θ_{jgp}.

Our main goal is to identify the intercept and slope parameters for every city $j \in \{1, ..., J\}$ and migration group $g \in \{\mathcal{N}, \mathcal{I}\}$. That is, we aim to identify the vector $\theta_j = (\alpha_{j\mathcal{N}}, \eta_{j\mathcal{N}}, \alpha_{j\mathcal{I}}, \eta_{j\mathcal{I}})'$, which provides a complete description of the causal effects of location j and how these effects vary with parent income and immigrant status.

3.2 Identification Strategy and Research Design

The ideal experiment would randomly send children to different places when the children were at different ages. Absent such an experiment, we turn to exploit a quasi-experimental design on the entire population, following Chetty and Hendren (2018b). We identify location effects by exploiting the variation in children's exposure time to different cities during childhood due to household moves at different ages. Our strategy combines variation in the timing of moves across locations within Israel with variation in the age at which children migrated to Israel from the former Soviet Union. To build intuition, consider the following example. Among all the native-born families that moved between city j to city l and are of the same income level,[4] some children arrived at younger ages, while some arrived at older ages. Then, if among that narrow group, the moving decision is unrelated to the child's age at the move, by comparing the outcomes of children who spent different time spans in each city, one can infer the effect of growing up in city j compared to city l: $\theta_{j\mathcal{N}p} - \theta_{l\mathcal{N}p}$.

[3] In Appendix Figure A.1, we present the relationship between children's income rank at ages 28-30 and parental income rank both in the full population and within a few selected cities, providing evidence that the relationship between children and parental income rank is approximately linear in Israel as well.

[4] The exact variable definitions, including parents' income, is described in Section 4.

Building on this logic, we exploit the variation in children's exposure time to different locations in Israel among all the families that experienced up to two moves when the child was young. Among immigrants, we include two groups. The first group includes the families that moved once to Israel when the child was at age a_i, settled in the city j, and stayed there until the child grew up. Among these families, the exposure variable is $E_{ij} = 18 - a_i$ for the first city of residence j, and zero otherwise. The second group is comprised of immigrants who moved twice. First, they immigrated to Israel when the child was at age a_i, settled in city $d(i)$, and then moved to city $d_2(i)$ when the child was at age a_{2i}. Among these families, exposure is given by:

$$E_{ij} = \begin{cases} a_{2i} - a_i, & \text{if } j = d(i) \\ 18 - a_{2i}, & \text{if } j = d_2(i) \\ 0, & \text{otherwise.} \end{cases}$$

Similarly, among natives, our analysis includes families who moved once or twice between cities in Israel before the child turned 18 years old. We denote m_i as the child's age at the first move from origin city $o(i)$ to destination city $d(i)$, and m_{2i} as the child's age at the second move to destination city $d_2(i)$, which equals zero if that child moved only once during childhood. Therefore, their exposure variable is given by

$$E_{ij} = \begin{cases} m_i, & \text{if } j = o(i) \\ m_{i2} - m_i, & \text{if } j = d(i) \\ 18 - m_{2i}, & \text{if } j = d_2(i) \\ 0, & \text{otherwise.} \end{cases}$$

Our approach diverges from the strategy pursued by Chetty and Hendren (2018a,b) by including not just one-time movers but also those who move twice. The main advantage of this approach is that it expands the sample size of our analysis.[5] In Appendix Section C, we explore the migration patterns of immigrants and natives between those who moved once and twice and between the first and second move. Surprisingly, we find that the distribution

[5]Adding those who moved twice increased the immigrant sample size by 23%, and the native-born sample size by 13%. For more detail, see Appendix Section C.

of origin-destination migration patterns in the first and second moves are pretty similar.

Under the model in Equation (1), we can conduct all the comparisons within families with the same sequence of location choices and parental income by estimating the following specification:

$$Y_i = \sum_{g' \in \{\mathcal{N},\mathcal{I}\}} \sum_{j=2}^{J} \left(\underbrace{(\alpha_{jg'} + \eta_{jg'} p(i))}_{\theta_{jg'p}} E_{ij} + x_i' \gamma_{g'} \right) \mathbb{1}\{g(i) = g'\} + \epsilon_i, \tag{3}$$

where x_i includes fixed effects for sequences of location choices and parental income rank, and $\epsilon_i = \xi_i(E_i) - \overline{\xi_i(E_i)}_{o,d,d_2,p}$ captures the variation in other human capital inputs within groups of families with the same migration patterns and income. By including the sequence of location choices fixed-effects, θ_j is identified only from variation in the timing of moves rather than variation between families that moved between different places.

Let $E = (E_1', .., E_N')'$ denote the $N \times (\mathcal{J} - 1)$ design matrix of location exposures, $X = (x_1', .., x_N')'$ the $N \times K$ design matrix of the fixed effects and controls where K is the number of sequences of location choices and family income, and $\epsilon = (\epsilon_1, ..., \epsilon_N)$ the $N \times 1$ vector of the error term. For OLS regression of Equation (3) to identify the location effect parameters of interest, we need the following strict exogeneity assumption:

Assumption A1 *(Strict exogeneity)*

$$E[\epsilon | E, X] = 0.$$

Assumption A1 imposes important restrictions on the economic environment. It requires that among children with the same set of childhood places and parental income, the time spent at each location is not systematically correlated with unobserved inputs that determine human capital. Importantly, it does not preclude systemic spatial sorting that correlates with the location effects of the origin and destination locations. For example, we find in Section 6 that immigrants are more likely to reside in cities with high long-run effects on children's income in adulthood.

Although Assumption A1 is a strong restriction, evidence from the US suggests that it is consistent with the data from both quasi-experimental (Chetty and Hendren, 2018a,b) and experimental (Chetty et al., 2016) settings. Moreover, in Appendix Section D, we provide further supporting evidence that this assumption is also consistent with our setting.

Empirical Implementation - Our empirical analysis includes a few modifications to Equation (3) to account for the specific features of our data. In particular, in Equation (3) x_i includes fixed effects for sequences of location choices at the $o(i)$-$d(i)$-$d_2(i)$ level for native-born children and at the $d(i)$-$d_2(i)$ by birth cohort for immigrants.[6] Because we measure children's outcomes at a fixed age and, therefore, in different calendar years, we add the birth cohort fixed effects to account for fluctuations in labor market conditions over time. Lastly, x_i includes year of birth fixed effects interacted with parental income rank, where we control for the sequences of location fixed effects and parental income in an additively separable way due to the sample size restriction we face.

We estimate Equation (3) among children whose families moved between cities in Israel or immigrated to Israel before the children turned 18 years old. In Equation (3), location effects are identified only in relative terms. Therefore, in our analysis, we set the base-level location of immigrants to be the former Soviet Union and the base-level location of native-born children to Jerusalem.

Estimation results in two vectors for every immigration group $g \in \{\mathcal{I}, \mathcal{N}\}$: one for location effect intercepts, $\hat{\alpha}_g = (\hat{\alpha}_{1g}, ..., \hat{\alpha}_{Jg})'$, and another for parental income rank slopes, $\hat{\eta}_g = (\hat{\eta}_{1g}, ..., \hat{\eta}_{Jg})'$, and their corresponding variance-covariance matrix, which is clustered by family id. The full estimated location effects vector is represented by the stacked vector $\hat{\theta} = (\hat{\alpha}'_{\mathcal{I}}, \hat{\eta}'_{\mathcal{I}}, \hat{\alpha}'_{\mathcal{N}}, \hat{\eta}'_{\mathcal{N}})'$, and its corresponding variance is represented by the matrix Σ. We are interested in studying the joint distribution of θ and measuring the heterogeneity in location effects across immigration groups.

3.3 Variance Components

Having estimated $\hat{\theta}$, our central objective is to study the heterogeneity in location effects both across cities and within cities by immigration group and parental income. We measure the heterogeneity across and within cities by studying the variance-covariance matrix of θ_j, denoted by Ω. For every group g, the diagonal elements of Ω give the variance of the elements

[6] For immigrants, we interact the sequence of location choices' fixed effects with the child's year of birth to account for the potential correlation between parents' cohorts and children's age at arrival. We thereby compare immigrant families that moved at different years within cohorts.

of θ_j. For example the variance of α_{jg} is

$$\sigma^2_{\alpha g} = \sum_{j=1}^{J} \frac{n_j}{N}(\alpha_{jg} - \sum_{j=1}^{J} \frac{n_j}{N}\alpha_{jg})^2 \tag{4}$$

where n_j is the number of children residing in city j during childhood for at least one year, and $N = \sum_{j=1}^{J} n_j$. The off-diagonal elements of Ω are the covariances of elements in θ_j with either the other group's parameter or the within-group relationship between the slope and the intercept.

We observe only noisy estimates of the location effects $\hat{\theta}_j$, rather than the location effects themselves, θ_j. Therefore, the sample variance, $\sum_{j=1}^{J} \frac{n_j}{N}(\hat{\alpha}_{jg} - \sum_{j=1}^{J} \frac{n_j}{N}\hat{\alpha}_{jg})^2$, of α_{jg}, or each of the other elements in θ_j, is over-dispersed. The standard approach to bias-correct the estimate of Equation (4) is to subtract from the sample variance the mean squared of the standard errors (Chetty et al., 2014a; Chetty and Hendren, 2018b; Rose et al., 2022; Kline et al., 2022). As detailed in Appendix Section F, we use a variant of that estimator, which accounts for the correlation of the $\hat{\theta}_j$ across different j arising from our estimation procedure.

4 Data

We use administrative data assembled by the Israeli Central Bureau of Statistics (CBS). The data covers the entire population of registered Israeli citizens for the birth cohorts of 1950-1995 and their parents, totaling an average of roughly 100 thousand individuals per birth cohort. The data comprise four primary sources: tax records, education records, the population registry, and the censuses of 1995 and 1983. From the Tax Authority, the data include employer-employee and self-employment tax records at a yearly level for the years 1995-2019. From the Ministry of Education, we match each child's school identifiers and city, matriculation scores and subjects, post-secondary degree attainment, and major. From the civil registry records, we match detailed information on demographics, including gender, year of birth, date of immigration and country of origin of children, parents, and grandparents, identifiers of siblings, an identifier of the spouse, if applicable, year of birth for every child, and, importantly for this project, the place of residence and school at the locality level (such as city, town, or village) on a yearly basis. Lastly, the 1983 and 1995 censuses, including the location of residence at a city level, education records, and earnings, are also included in our data. For the rest of this section, we describe the sample definition and the main variables we use. Further details are in Appendix B.

4.1 Sample Selection and Variable Definitions

The main sample consists of all children born in the years 1980-1995. Each child in our sample has a unique identifier for both parents. Using the location of residence of both the child and the parents, we define the primary parent to be the one who shares an address with the child for the majority of the years. If a location value is missing for a certain year, we fill the location of residence with the child's school location only if the school is in the same location as the child's location of residence in year $t-1$.[7] We enrich the location data with the city information available from the 1995 census. Specifically, we use the answers to two questions: "When did you move to your current city?" and "Where did you live five years ago?". Using these variables, we can construct location information starting from 1995 and, for a subset, from 1990.[8] For the rest of this paper, our unit of location is a city or regional council[9], which serve as the formal identifiers for cities in Israel and represent the units of local government.[10]

For every parent in the sample, we construct the following variables: Parents' income, which is the total gross income at the household level, measured in 2014 Israeli Shekels (1 ILS $\sim$ \$0.28). In years when the family has no recorded earnings, the family income is coded as zero. To derive an approximation of parents' resources during childhood, we calculate the average earnings over the years 1995-2016. This time frame is selected to balance between potential attenuation biases that may arise from measuring parental income over too short a period and the risk of doing so too late in life when income tends to be more volatile (Mazumder, 2005). We exclude families with less than four years of earnings in this time period, which accounts for 1.5% of the parents. Finally, we work with a parents' percentile rank variable, defined as the parent income rank in the national population that satisfies the restriction of having at least four years of earnings in 1995-2016. To account for the

[7]Therefore, we do not identify moves resulting from the school location being different than the location of residence.

[8]Response rate to the questions in 1995 is around 20%.

[9]A regional-council locality is a group of small localities such as small towns or *Kibutzim* that are geographically close to each other and share the same local governing council

[10]Although we do observe a smaller unit of location, equivalent to a census tract in the US, our data agreement usage restricts us to estimate location effects only at the city level or bigger geographic unit.

unbalanced structure of the child's age at which parents have earnings, we calculate each income rank within children's cohorts, therefore comparing parents' earnings for children at the same ages.

As our main outcome for children, we use the income at age 28 and calculate the percentile rank in the general population within the child's cohort to account for differences in calendar year differences in labor market fluctuations. Additionally, using the administrative records on post-secondary degree attainment, we measure for every child whether they attain a bachelor's degree by age 27.

We work with two primary populations. The first group consists of immigrants from the former Soviet Union (FSU) who arrived in Israel between 1989 and 2000. We identify the children of the immigrants based on their parents' birth country and year of immigration.[11] For each immigrant child, we calculate a_i, the age of the child when the family immigrated to Israel. We then designate the first city or regional council of residence as their initial destination location and record any other cities where the family lived during the child's childhood had they moved.

The second group in our analysis is the "native-born", which includes all non-Arab individuals born in Israel (including families from older immigration waves).[12] Similarly to the immigrants, for every family, we record all the cities in which the families lived during their children's childhood. In some cases, we refer to the families residing in a single location throughout the child's childhood as permanent residents or stayers and the subset of families that are not permanent residents as movers.

Lastly, our objective is to measure the childhood location effect separately for every city and immigration group. We restrict attention to cities with at least 100 individuals in every group. As such, these requirements narrow our analysis to a list of 98 cities and regional councils out of 253, for which we have sufficient observation for both immigrants and locals.

[11] Around 10% of the immigrants during this period had a missing country of origin before immigration. In such cases, we have classified them as FSU immigrants as well, since immigrants from the former Soviet Union accounted for 90% of the arrivals during these years (See Figure 1).

[12] Approximately 20% of the Israeli population are Arab citizens; however, since Jews and Arabs in Israel are geographically segregated, where over 70% of the Israeli Arabs live in cities and villages that are 100% Arabs, we have very little overlap between the two groups that would restrict our ability to compare them.

4.2 Summary Statistics and the Immigrant-Native Income Gap

Table 1 presents the number of children and mean income of parents and children in Israeli Shekels ($\$1 \sim 3.4$ ILS) of Soviet immigrants and natives for all the immigrant families that either moved to Israel and stayed in the same city or immigrated to Israel and then moved between cities in Israel before their child turned 18. Among native-born families, the table displays the characteristics of families that either moved once or twice between cities before the children turned 18. For a more detailed comparison between the families that moved once and twice, see Appendix Section C.

The selected 98 cities are the largest cities and regional councils in Israel and, therefore, represent most of the Israeli citizens. 88% of the immigrants and 81% of the native-born families that move between cities in Israel are in our sample. Parental income in our selected sample of cities is slightly lower for both groups compared to the families in the full city sample, while children's income is slightly higher. Notably, immigrant parental income is 55% that of the natives, echoing the results in Cohen-Goldner and Paserman (2011), Goldner et al. (2012), and Arellano-Bover and San (2023) of a large immigrant-native wage gap. Arellano-Bover and San (2023) find that the adult wage gap closed only after 27-29 years after arrival in Israel. However, the second generation closed most of that gap with immigrant children's income at age 28, attaining 95-96% of native-born children.

In Appendix Figure A.2, we plot the geographic distribution of immigrants across Israel, both as shares of the immigrant population across Israel (Panel (a)) and as a share of the local population within each locality (Panel (b)). City names of the top 10 shares in each map are printed on the map. As expected, the biggest Israeli cities, such as Haifa, Tel-Aviv, Jerusalem, and Be'er-Sheva, absorbed the largest proportion of immigrants. Panel (b) highlights that immigrants did not settle only in the big urban cities but throughout the country and accounted for a substantial share of the residents in multiple cities.

5 Estimates of Location Effects

5.1 Estimation Results

Our multidimensional model for childhood location effects allows for heterogeneity both across cities and within cities by parental income and family immigration status. We start

Table 1: Descriptive statistics

	All cities		98 cities	
	Immigrants (1)	Native-born (2)	Immigrants (3)	Native-born (4)
(A) Children				
Income age 28	67,108	70,741	68,191	71,701
Rank age 28	52.5	53.6	53.2	54.2
(B) Parents				
Parents' income	131,670	235,981	129,997	233,095
Rank parents	45.7	63.3	44.8	63.1
Num. of children	156,269	116,572	138,664	95,500

Note: This table presents the mean children's income and income rank at age 28 and the mean parental income and parental income rank between the years 1995-2016 among immigrants and natives that are in our main sample of movers. All income variables are measured in Israeli Shekels (1 US \$ $\approx$ 3.4 ILS). Among immigrants, the sample includes all the immigrants that either arrived in Israel and stayed in the same city or arrived in Israel and then moved again between cities in Israel before the children turned 18. Among natives, the sample includes all the families that moved either once or twice between cities in Israel before the children turned 18. Columns 1-2 present the statistics for all the families, and columns 3-4 present the statistics for families in our selected sample of 98 cities and regional councils. Panel (A) displays children's mean income and mean income rank at age 28. Panel (B) displays parental income and parental income rank at the national distribution.

by exploring the extent to which the effect of childhood locations varies within immigration groups and across cities and parental income and then study the within city heterogeneity across immigrants and natives.

5.1.1 Across City Heterogeneity

Table 2 presents estimates of the distribution of the causal location effects. Panel (i) reports the mean and standard deviation of α_{jg} and η_{jg} for immigrants and natives. As noted in Section 3.2, the cardinal value of location effects is not identified. For natives, the location effects measure the effect of spending one more year in city j compared to Jerusalem, while for immigrants, the location effects measure the effect of spending one more year in city j compared to one more year in the former Soviet Union. To summarize the full one-year effect of each city, panel (ii) presents the same statistics for $\theta_{jpg} = \alpha_{jg} + \eta_{jg} \times p$, for the

$p = 25$ and $p = 75$, which we refer to as the location effects of low and high-income families, respectively. Columns (1) to (3) display the main statistics for all the cities in Israel that satisfy the sample restrictions separately for each immigration group, and columns (4) to (6) present them for the set of overlapping locations for which we estimate location effects for both immigrants and natives. Column (7) reports a Wald test statistic and corresponding p-value for the null hypothesis of no location effect heterogeneity across the overlap cities.

The causal effects of neighborhoods vary substantially across income groups and by income status. There are a total of 150 cities and regional councils that satisfy the sample restriction for immigrants and 99 cities and regional councils that satisfy the sample restriction for natives. The average intercept α_{jg} of natives is 0.25, implying that an extra year spent in the average city rather than Jerusalem boosts age-28 income for native-born children at the lowest income rank by 0.25 ranks. The corresponding estimate for immigrant children is 0.09 (SE=0.032) income ranks relative to staying one year in the Soviet Union.[13] The slope on parental income rank $\eta_{j\mathcal{N}}$ captures heterogeneity in the effect of location j by parent income relative to Jerusalem for natives and relative to the USSR for immigrants. Because the estimates are relative to a base-level location, we can get negative values of η_{jg}.[14]

Finally, panel (ii) summarizes the distribution of location effects separately for low- and high-income families, separately by immigrant status. In line with the finding in panel (i), the mean location effects for families at the 25th percentile are positive and statistically distinguishable from zero. For every year spent in the average city, low-income immigrant (native) earnings rank at age 28 increases by 0.123 (0.180) compared to one more year in the USSR (Jerusalem). This rank increase is equivalent to 307 (264) Israeli Shekels increase, which amounts to 90 (77) US dollars. However, comparing the mean effects of natives and immigrants reveals heterogeneity in location effects with respect to parental income. While the effect of the average city on a child's income for low-income families is higher than the

[13]This result suggests that spending one more year in the average Israeli city increases child income among immigrant children. As discussed in Appendix Section G and at the end of this section, we provide further analysis and test whether longer exposure time to all the cities in Israel generates income gains compared to staying in the USSR.

[14]While the level of the intercept θ_{jg} is not identified, we test whether the slope η_{jg} is always positive as one would expect. In our fixed-effect model, we control for the interaction of parents' rank and children's birth cohorts. To test whether η_{jg} is always positive, we average the estimated coefficients of parental income rank and child birth cohorts and add that number to $\hat{\eta}_{jg}$. The Bai et al. (2022) testing procedure suggests we can reject the null that η_{jg} is always negative, with p-values of > 0.01, and cannot reject the null that it is always positive, with p-values of > 0.99.

Table 2: Variation in location effects on adult income rank at age 28

	All cities			Overlap cities			
	# of cities	Mean	Std.	# of cities	Mean	Std.	χ^2 test $H_0 : \theta_j = \theta_1 \forall j$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(i) By α and η							
Natives							
Cons. (α)	150	0.253	0.222	98	0.254	0.210	142.4
		(0.095)	(0.050)		(0.101)	(0.053)	[0.0023]
Rank-parents (η)	150	-0.003	0.003	98	-0.003	0.002	141.5
		(0.001)	(0.001)		(0.001)	(0.001)	[0.0027]
Immigrants							
Cons. (α)	99	0.093	0.228	98	0.120	0.228	187.1
		(0.032)	(0.037)		(0.034)	(0.042)	[0.0000]
Rank-parents (η)	99	0.001	0.002	98	0.001	0.002	188.0
		(0.001)	(0.000)		(0.001)	(0.000)	[0.0000]
(ii) Total city effect							
Natives							
θ_{25}	150	0.180	0.175	98	0.184	0.173	138.4
		(0.086)	(0.043)		(0.091)	(0.044)	[0.0045]
θ_{75}	150	0.036	0.159	98	0.043	0.155	139.5
		(0.085)	(0.043)		(0.090)	(0.043)	[0.0037]
Immigrants							
θ_{25}	99	0.123	0.239	98	0.143	0.246	212.8
		(0.026)	(0.034)		(0.028)	(0.037)	[0.0000]
θ_{75}	99	0.183	0.295	98	0.190	0.313	259.4
		(0.031)	(0.036)		(0.032)	(0.038)	[0.0000]

Note: This table presents estimates of the distribution of causal effects of Israeli cities on income rank at age 28, separately for immigrant and native children. Columns (1)-(3) show estimates for all available cities, while columns (4)-(6) display estimates for cities with sufficient samples to estimate effects for both immigrants and natives. Estimates come from OLS regressions of child income rank on years of exposure to each location and interactions of years of exposure with parent income rank controlling for location sequences fixed effects and birth-cohort fixed effects interacted with parents' income rank. Panel (a) reports estimates of the distributions of location-specific intercepts (α) and slope coefficients on parent income rank (η). Columns (2)/(5) show the mean of each estimated parameter, and columns (3)/(6) show standard deviations, computed as the square root of the standard deviation of the bias-corrected variance of parameters across locations. Panel (b) displays corresponding distributions of location effects for children at the 25th and 75th percentiles of parent income, computed as the sum of the location intercept and the parent income slope multiplied by the relevant percentile. Column (7) shows test statistics and p-values from chi-squared tests of the null hypothesis that all locations are identical. Standard errors for all variance estimators are based on the asymptotic variance, assuming the location effects are drawn from a normal distribution.

effect of Jerusalem, that effect is as good as Jerusalem for high-income families.

Column (3) presents the standard deviation of the city-level variance, which is the square root of the unbiased variance. The reported estimates of the variances of α_{jg}, η_{jg}, θ_{jg25}, and θ_{jg75} imply substantial across-city variation in location effects in Israel among different immigration and income groups. The standard deviation of the total location effects for families at the 25th (75th) percentile is 0.175 (0.159) for natives and 0.239 (0.295) for immigrants. These estimates are of a similar magnitude to the standard deviation of location effects in children's income rank found in the US across counties (Chetty and Hendren, 2018b). Interestingly, we find that locations are more consequential for immigrants than for natives. For every income level, the variance of the location effects of immigrants is greater than the variance of natives. Among low-income families, the variance of the location effects of natives is 54% of that of immigrants, while among high-income families, it is 30% that of immigrants.

Moving at birth to a city with one standard deviation higher location effect for natives (immigrants) increases children's income rank at age 28 by $0.17 \times 18 = 3.06$ ($0.23 \times 18 = 4.14$) points. To assess the magnitude of location effects, we rescale the one-year location effects to money value. By regressing separately by parental income rank children's income on their rank among children who spent all their childhood in one city, we find that every percentile rank increase in children's income yields 1,530 additional Shekels ($\approx$ \$450) for families at the 25th percentile and 1,689 Shekels ($\approx$ \$490) for families in the 75th percentile. Therefore, one standard deviation better city at birth increases the income of native-born (immigrant) children from the 25th percentile by 4,681 (6,992) ILS, or \$1370 (\$2056), which is 8% (9%) of the mean income of children with parents with below median income.[15] For comparison, the return to matriculation certificate in Israel is 13% (Angrist and Lavy, 2009). Thus, moving at birth to one standard deviation better city yields 61-70% of the gains from a matriculation certificate attainment.[16]

Columns (5)-(6) display the same statistics among the 98 cities and regional councils for which we estimate location effects both for immigrants and natives that satisfy our sample restriction. For the remainder of the paper, we use this sample to study location effects in Israel and how they vary between immigrants and natives. As reassurance that this is not a special set of cities, we note that the estimates of the first two moments presented in columns (2) and (3) are not qualitatively different from the estimates in columns (5) and (6). Finally, in column (7), we present the χ^2 test statistic and corresponding p-value for the null hypothesis of no location effect heterogeneity across the overlapped cities. For all the city-level parameters, we reject the null of no location effect heterogeneity at conventional

[15] The average income at age 28 of immigrant children from a below-median income family is 59,670 Shekels, and 77,111 above the mean. The corresponding values for natives are 57,884 Shekels for below-median families and 73,448 above.

[16] To secure a matriculation certificate, students must pass a series of standardized national exams administered during the final two years of high school. The attainment of a matriculation certificate is a critical factor in children's future labor market outcomes, as nearly all post-secondary institutions mandate it for admission.

significance levels.

The evidence in Table 2 suggests that among immigrant children, spending one more year in the average Israeli city increases their income rank in adulthood. This raises the follow-up question: is this true not only for the average city but for all cities in Israel? In Appendix Section G, we further investigate this question. Our findings suggest that, on average, immigrating at a younger age had a positive effect on children's income in adulthood, as found similarly in Alexander and Ward (2018). Nevertheless, we also show that there are places that induce negative effects, therefore extending the finding's mean effects. This suggests that the determinants of the returns for early immigration also depend on the local environment within the destination country where the immigrant lives.

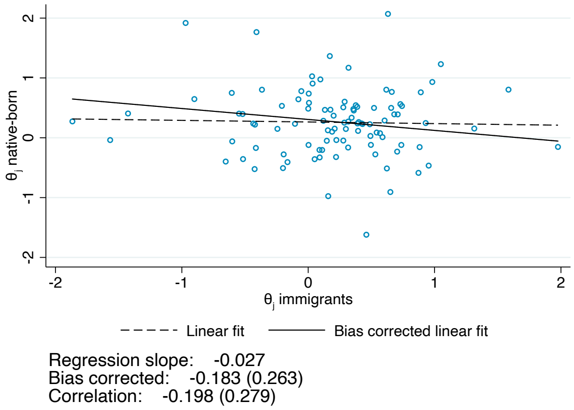
5.1.2 Immigrant-Native Differences in Childhood Location Effects

Location effects vary substantially between native and immigrant children at the same income level. This is revealed in Figure 2, which plots scatter plots and observation weighted regression lines of the estimates of effects for natives against the corresponding effects for immigrants, separately by income group. Figure 2a displays the relationship between immigrants' and natives' intercepts α_{jg}, i.e., between the location effects on families from the lowest income percentile, Figure 2b displays the relationship between the slopes η_{jg}. i.e., between the city returns to parental income, and Figures 2c and 2d display that relationship for the total one-year location effects for families at the 25th and 75th percentiles of the national income distribution. Dashed lines are the naive attenuated regression lines, while solid lines display the biased corrected regression line, with slopes estimated as the ratio between the covariance and the bias-corrected variance of immigrants' locations effects. Table 3 presents the corresponding estimates, together with the mean and standard deviation of the within-city immigrant-native location effects gaps.[17]

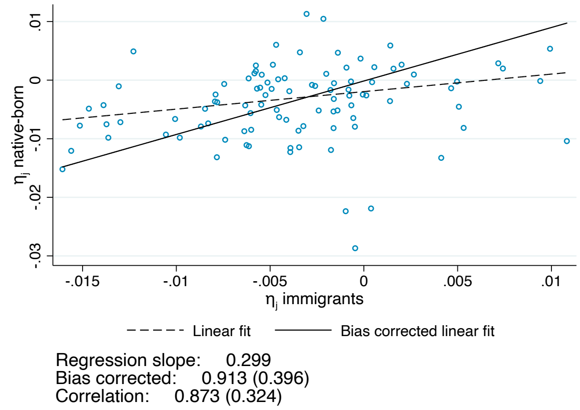
The scatter plot and regression lines of the intercept in Figure 2a imply substantial heterogeneity between immigrants and natives with the lowest parental income. Places that benefit low-income immigrants are not necessarily places that benefit low-income natives, where the point estimate of the bias-corrected correlation is actually negative ($corr = -0.20$). In contrast, Figure 2b suggests much less heterogeneity in location effects as parental income increases. Places with high returns to increasing parental income for immigrants are likely to be places

[17]The full correlation matrix of $(\alpha_{j\mathcal{N}p}, \eta_{j\mathcal{N}p}, \alpha_{j\mathcal{I}p}, \eta_{j\mathcal{I}p})'$ is reported in Appendix Table A.1

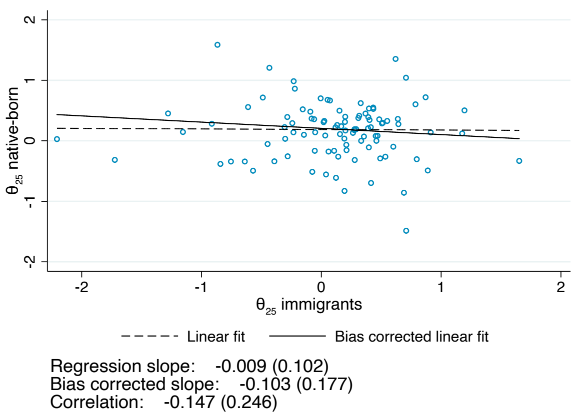
Figure 2: The relationship between location effects for immigrants and natives



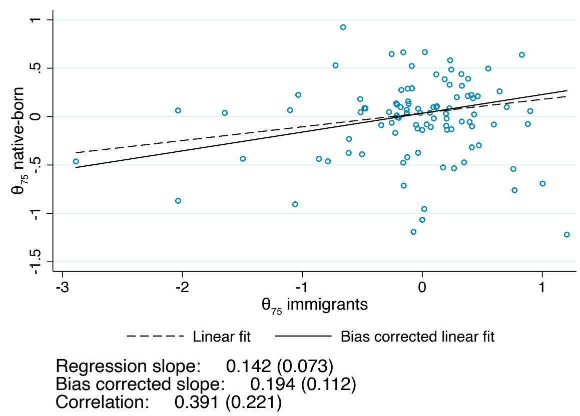
(a) α (intercept)



(b) η (slope)



(c) Total effect, θ_{25}



(d) Total effect, θ_{75}

Note: These figures display the scatter plots regression lines of immigrants' and natives' location effects. Panel (a) plots the estimated intercepts α_{jg} of immigrants and natives, panel (b) plots the estimated slopes η_{jg} of immigrants and natives, panel (c) plots the total one-year location effect for families at the 25th percentile of the income distribution, and panel (d) plots the total one-year location effect for families at the 75th percentile of the income distribution. The dashed line is the naive regression line, and the solid line is the bias-corrected regression line using the slope from Table 3. All regression lines are weighted by the number of observations.

with high returns to increasing parental income for natives. Combining these two facts, we find in Figure 2c no relationship between the location effects of immigrants and the location effects of native families who are at the 25th percentile of the national income distribution, while for families at the 75th percentile (Figure 2d), the location effects of immigrants and natives are strongly correlated. A correlation of 0.40 between the location effects of high-income immigrants and natives implies that places with one standard deviation higher effects for high-income immigrants are associated with almost half of a standard deviation higher location effects for natives.[18]

The standard errors on the correlation coefficients, calculated via the delta method, suggest that these correlations are imprecisely estimated. At the same time, the correlation is a highly nonlinear function for which the delta method approximation may be inaccurate. Therefore, we report in the square brackets of column 3 of Table 3 the bootstrapped equal-tailed 90% confidence intervals under the assumption that location effects are normally distributed. We use these confidence intervals to perform the one-sided test for each correlation coefficient being different from 1. Among low-income families, either at the bottom of the income distribution or at the 25th percentile, we can decisively reject any positive correlations stronger than 0.3. In contrast, for η_{jg}, the return to parental income, we find a correlation of 0.87 where we cannot reject the null that the correlation coefficient equals 1.

In the last three columns of Table 3, we present the mean and standard deviation of the difference between immigrants' and natives' location effects and the p-value for the test that the within-city immigrant-native differences are equal across all cities. First, column (5) reveals substantial heterogeneity in the city-level gap between the location effects of immigrants and natives. The standard deviation of the city-level gap is 0.32 and 0.29 for families at the 25th and 75th percentile of the national income distribution, which is 17-76% higher than the standard deviation of the effects themselves. This finding implies that moving at birth to a city with one standard deviation higher differences between immigrants and natives implies moving to a city that increases the adulthood income of one group by 8,598 ILS ($\approx \$2,623$) more than the other group, which is more than 14% of the mean income at age 28 for children from a below-median income family. As another benchmark, Appendix Figure A.5 shows that in the 98 cities in our analysis, the immigrant-native upward mobility gap among families from the 25th percentile is between 1 and 3.5, depending on the age at the time of income measurement. Note that 2 rank points are the average causal gap arising from spending seven years in a city with one standard deviation higher location effects gaps.

Column (6) presents the p-value for the null hypothesis of no within-city differences in

[18] The full correlation matrix in Appendix Table A.1 shows that the point estimate of the correlation between the intercept and the slopes is negative and imprecisely estimated. This negative correlation can be partially explained as a mechanical boundary correlation as the outcome is bounded between 0 and 100, and the level value of the slope η is positive (see footnote 14).

Table 3: Differences in location effects between immigrants and natives

	Covariance	Correlation	Implied OLS coefficent	Difference: Mean	Difference: Std.	Difference: χ^2 test $H_0: \theta_{j\mathcal{N}} - \theta_{j\mathcal{I}} = c\ \forall j$
	(1)	(2)	(3)	(4)	(5)	(6)
α	-0.010	-0.198	-0.183	-0.134	0.340	166.9
	(0.013)	(0.279)	(0.263)	(0.106)	(0.061)	[0.000]
		[-0.710, 0.290]	[-0.675, 0.235]			
η	0.000	0.873	0.913	0.004	0.001	105.5
	(0.000)	(0.324)	(0.396)	(0.001)	(0.002)	[0.285]
		[0.448, 1.405]	[0.531, 1.754]			
θ_{25}	-0.006	-0.147	-0.103	-0.040	0.321	177.9
	(0.010)	(0.246)	(0.177)	(0.095)	(0.051)	[0.0000]
		[-0.613, 0.293]	[-0.446, 0.213]			
θ_{75}	0.019	0.391	0.194	0.148	0.290	162.0
	(0.011)	(0.221)	(0.112)	(0.095)	(0.051)	[0.0001]
		[0.014, 0.817]	[-0.011, 0.417]			

Note: This table reports the relationship between the location effects of immigrants and the location effects of natives and a test for within-city heterogeneity. Column (1) presents the covariance between the location effects of immigrants and natives, column (2) presents the bias-corrected correlation, which is the covariance divided by the standard deviation of immigrants times the standard deviation of locals, and column (3) presents the implied OLS coefficient, which is the covariance divided by the variance of immigrants. Column (4) presents the mean within-city gap between immigrants and natives, column (5) presents the standard deviation of the within-city gap, and column (6) presents test statistics and p-values from chi-squared tests of the null that location effects don't vary within cities. Location effect estimates come from OLS regressions of child income rank on years of exposure to each location and interactions of years of exposure with parent income rank controlling for location sequences fixed effects and birth-cohort fixed effects interacted with parents' income rank. The first row reports estimates of the location-specific intercepts (α), the second row reports the estimates of the slope coefficients on parent income rank (η), and the last two rows report location effects for children at the 25th and 75th percentiles of parent income distribution, computed as the sum of the location intercept and the parent income slope multiplied by the relevant percentile. Standard errors of the variance and covariances are based on the asymptotic variance, assuming location effects are drawn from a normal distribution. Standard errors of the correlations and OLS slopes are calculated using the delta method. Square brackets display parametric bootstrapped equal-tailed confidence intervals.

location effects. In line with our findings, we can decisively reject the null of no within city heterogeneity, except for the slope coefficients η_{jg}.

6 Predictors of Location Effects

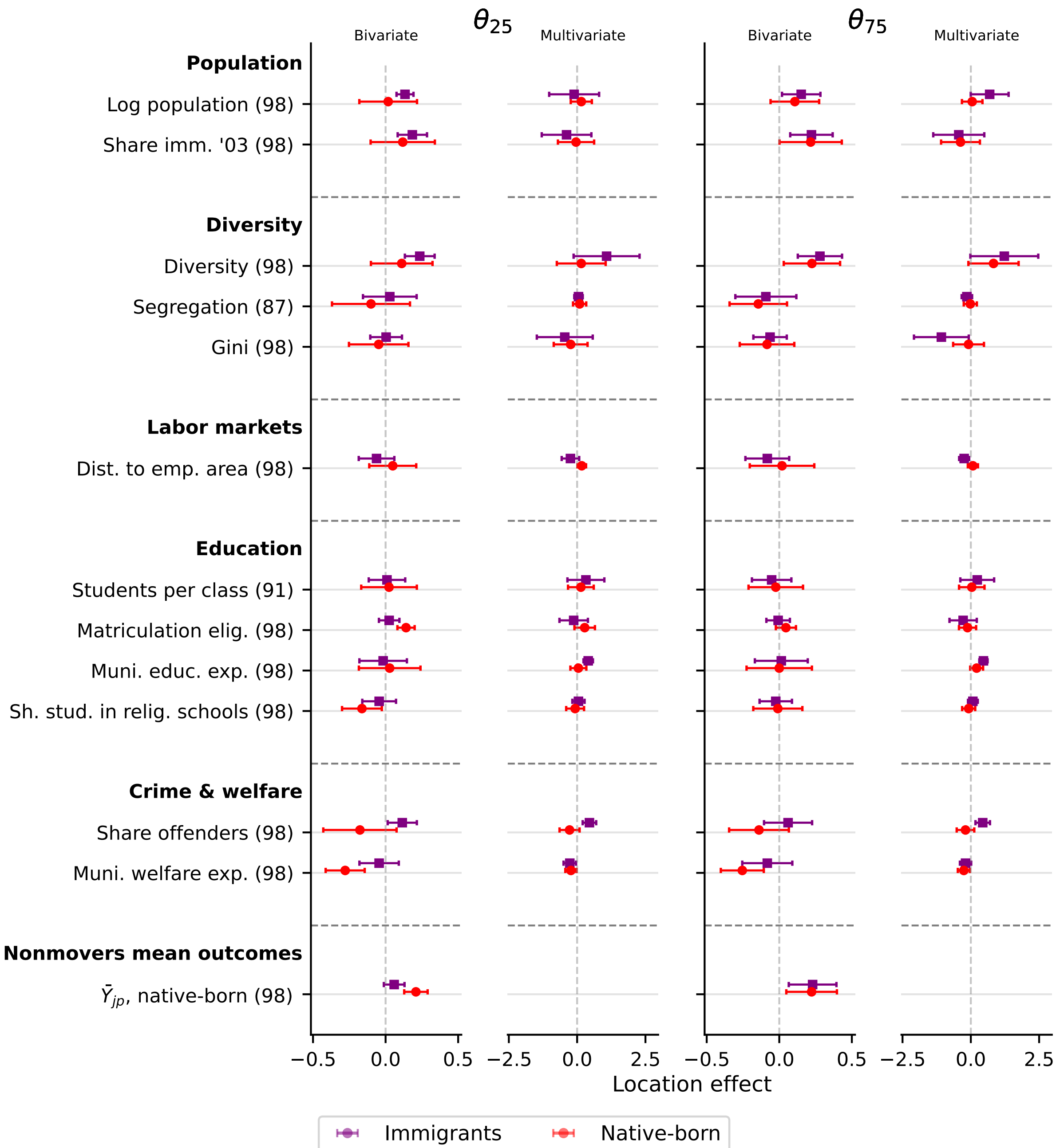
Next, we explore the characteristics of cities with high long-run effects on children's income by estimating the linear relationship between effects and characteristics at the city level. Throughout this section, within each group, immigrants and natives, we demean the effects and the characteristics and divide them by the sample standard deviation. Therefore, the estimated coefficients are correlations measured by units of standard deviations. For most locality-level characteristics, we rely on data from the early 2000s collected from various sources. Detailed definitions of the variables and information about their sources can be found in Appendix Section B.3.

Figure 3 plots the coefficients from the feasible generalized least squares regression of the total one-year effect of immigrants and natives from the 25th and 75th percentile of the national income distribution on city characteristics, reweighting the inverse of the Cholesky decomposition matrix of Σ, the variance of $\hat{\theta}$. The relationship with α_{jg} and η_{jg}, which describes the location effects on families at the bottom of the income distribution, and the return to parental income is presented in Appendix Figure A.6.

Population and diversity - The first relationship we explore in Figure 3 is between population size and location effects. We find that larger cities are associated with larger long-run causal effects on children of immigrants. Larger cities tend to enable large concentrations of ethnic and minority groups. Therefore, next, we explore the relationship between immigrant share and location effects. We find that cities that benefit low-income immigrant children tend to have a larger immigrant share, where, as explored in Appendix Figures A.7a and A.7c, this bivariate relationship is approximately quadratic.

The literature on the effects of geographic concentration of ethnic groups on their economic outcomes is mixed. The effects of ghettos on minority groups, mostly the black population in the US, have been argued to be negative and long-lasting (Massey and Denton, 1993; Cutler and Glaeser, 1997; Chyn et al., 2022a,b). In contrast, the evidence on refugees points to a more complex relationship. On the one hand, a handful of papers on immigrants and refugees find that enclave size positively affects the labor market outcomes of refugees, emphasizing the role of networks and social support (Edin et al., 2003; Beaman, 2012). Interestingly, in these papers, the shares of immigrants in the city were at most 10%. On the other hand, in a recent contribution of large Jewish enclaves in New York from the beginning of the 20th century, Abramitzky et al. (2020) show that Jewish immigrants who left the enclave experienced an increase in their earnings as well as their children's earnings. In that setting, the Jewish enclaves were huge, comprising over 60% of Jews. Combining this evidence with the descriptive relationship in our data raises the hypothesis that immigrants' prosperity and assimilation increase with city diversity rather than only group share. That is, immigrants benefit from a balance between sufficient support from their own community, on the one hand, and adequate exposure to the general population, on the other hand.

Figure 3: Relationship between location effects and city characteristics



Note: This figure plots the relationship between city-level covariates and the total location effect of yearly exposure for high- and low-income families whose income rank is at the 25th (left panel) and 75th percentile (right panel) of the income distribution. Each relationship is estimated with a feasible generalized least squares regression, reweighting the observations by the inverse of the Cholesky decomposition matrix of Σ, the variance of $\hat{\theta}$, and with the location effects as the outcomes. Covariates are standardized to have a mean of zero and a standard deviation of one in the sample. In each panel, the first column plots the coefficients from regressions of effects on each covariate alone, and the second column plots the coefficients of a multivariate regression with all the characteristics simultaneously. Bars indicate 95% confidence intervals based on robust standard errors. Appendix Section B.3 provides a complete description of covariates definitions. The number of cities in each regression is in parentheses. Cases with fewer localities than the full sample (98) are due to missing values, or in the case of segregation, where values cannot be calculated for cities that do not have sub-areas (see Appendix Section H.1).

Building on this, we measure diversity using the log entropy index:

$$-\left(\pi_{j\mathcal{I}}\ln(\pi_{j\mathcal{I}}) + (1-\pi_{j\mathcal{I}})\ln(1-\pi_{j\mathcal{I}})\right)$$

which, with two groups, immigrants and non-immigrants, our diversity measure receives its maximum value when the immigrant share equals half and its lowest value when it is zero or one. We find a strong association between location effects and diversity for low-income immigrants, a relationship that remains significant in a multivariate regression that controls for all the characteristics being explored.

A positive correlation between group share and location effect could also reflect sorting where immigrants are more likely to locate in places that benefit their children in terms of long-run economic outcomes. In the US, Chetty and Hendren (2018a) suggest that low-income families are less likely to reside in areas with high long-run effects on children. Our findings suggest that this is also the case in Israel for low-income Native families but not for immigrant families.[19] Generally, our results call for more causal research to disentangle peer effects from sorting, with particular emphasis on the differences between immigrants and natives.

Next, we examine the correlations with measures of segregation and income inequality, finding no evidence for any observed correlation between these measures and location effects. For every city and regional council, we construct our segregation measure based on the Theil (1972) index, which measures the extent of within-city geographic segregation between immigrants and non-immigrants, and we measure income inequality by the Gini index in each place.

Labor markets - Several studies have posited that local labor market opportunities significantly influence children's future trajectory by facilitating access to such opportunities (Wilson, 1987; Garin and Rothbaum, 2022). To assess that, Figure 3 plots the coefficient of the relationship between the weighted average of the distance from the employment center, weighted by the number of workers in each center, and location effects. We find that there

[19] Abramitzky et al. (2019) find somewhat different results among immigrants in the US. They show that immigrants are more likely to reside in areas with high mobility rates. The mean area mobility rates probably reflect the mobility rates of natives, which suggests that immigrants live in places with high native-born long-run outcomes. Nevertheless, further evidence is required to compare their findings and ours as they don't estimate causal location effects but mean outcomes conditional on parental earnings.

is no clear relationship between the two. In an extended analysis, we additionally find that location effects are not correlated with employment rates and youth work. This finding is unsurprising given Israel's small geographical scale.[20]

Education - In the next panel on Figure 3, we study the relationship between location effects and education inputs and outputs. We find a no relationship between city-level education input and output measures location effects. In Section 7, we investigate the role of high schools in further detail.

Crime and welfare - Figure 3 shows that municipality welfare expenditure per capita is negatively predictive of native-born children's location effects of all family incomes, a relationship that also survives in the multivariate regression. In our data, these are our best proxies for city poverty rates.[21] Interestingly, municipality welfare expenditure per capital is *not* predictive of immigrant location effects. These associations further emphasize the heterogeneity in our data.

City-level mean child rank conditional on parental income - Previous research has emphasized that observable mean child rank conditional parental income rank is strongly predictive of location effects and suggested using these statistics for policy targeting (Chetty and Hendren, 2018b; Bergman et al., 2019). We illustrate here that due to the high heterogeneity in location effects, it is not predictive of the benefits of all the groups. In particular, the last row in Figure 3 presents the relationship between location effects and city-level native-born nonmovers mean income ranks conditional on parental income. We construct these estimates by running the following regression for each Israeli city only on the sample native-born children whose parents stayed in the same city throughout their childhood:

$$Y_i = a_{j(i)} + b_{j(i)} p(i) + u_i$$

[20] In the US, commuting zones typically serve as the geographical unit representing local labor markets. In contrast, Israel, being compact in size, approximates the size of merely one, or at most, two commuting zones. For example, Appendix Figure A.9 presents the number of workers across Israel's primary employment centers, as measured in the 2008 census. We can see that Israel has essentially one major employment center around the Tel-Aviv metropolitan area.

[21] Unfortunately, there are no official records of poverty rates at a city level.

where $p(i)$ is parental income rank and Y_i is child's income rank at age 28. We then generate the predicted upward mobility score of a child with family income rank p residing in city j as

$$\bar{Y}_{jp} = \hat{a}_j + \hat{b}_j p.$$

Finally, in the last panel of Figure 3, we present the relationship between each place effect for families at the 25th and 75th percentile in the income distribution and the corresponding $\bar{Y}_{jp}$ for $p = 25$ or 75. In line with our evidence for heterogeneity, the native-born nonmovers mean permanent residents outcomes are strongly predictive of natives' effect but hold very little predictive power of low-income immigrants' location effects. These measures serve as the main instrument for guiding housing voucher policy (Bergman et al., 2019). Their weak predictive power of low-income immigrant place effects hints at the potential risk that may arise when using them to guide policy. We further discuss this risk in Section 9.

To sum up, our analysis provides two new facts about the type of cities that benefit immigrants and natives. First, unlike natives, low-income immigrants do benefit from populated cities, especially if these cities are diverse and have sufficient representation of the different groups in the society. Second, previous literature emphasized the relationship between poverty rates and location effects. We find that in Israel, the poverty rate is only predictive of natives' location effects. The city poverty rate and similar proxies were the first measures the literature used for targeting housing policy (Katz et al., 2001), again hinting at the possible costs to immigrants from guiding policy based on such measures. In Section 9, we formally portray the caution policymakers should undertake when forming policies in the face of heterogeneity.

7 Robustness and Alternative Exercises

High school fixed-effect - We provide an approximation for the role of high schools in explaining the variability in location effects across cities and groups. We do so by comparing the variance component of location effects from Equation (3) and the variance components estimated in an almost identical regression that additionally controls for high school fixed effects.[22] To avoid dropping observations from our original location effects model, we group high schools with fewer than five observations into one category.[23] This model is identified by the fact that some cities have more than one school, and some schools accept children from several local surrounding cities.

Appendix Table A.2 presents the results where column (1) displays the variance components of our baseline model, column (2) displays the variance components of location effects with high school fixed effects, and column (3) displays the share of variance explained by high school fixed effects. We describe below the results for low-income families (panel (i)) and note that the evidence on high-income families is qualitatively similar.

We find that while the standard deviation of low-income native (immigrant) location effects is 0.17 (0.20) at baseline, it declines to 0.13 (0.17) when controlling for high-school fixed effect. Thus, high schools explain $1 - \frac{0.13^2}{0.17^2} = 45\%$ ($1 - \frac{0.17^2}{0.20^2} = 31\%$) of the variation in location effects. However, when we turn to study the heterogeneity between immigrants and natives, we find surprising results. The variance of the immigrant-native within city difference is three times larger, and although the correlation coefficient becomes much noisier, the point estimate is more negative. This suggests that disparities between immigrants and natives in location effects are not caused by schools. If anything, high schools in Israel act as equalizers. For high-income families, the drop in the correlation is even more striking, as, without the high-school fixed effects, the correlation was strongly positive.

Neighborhood reweighting - We find substantial heterogeneity between the location effects of immigrants and natives at the city and regional council levels. Nevertheless, one might worry that the true location effects vary substantially within cities at the neighborhood level. If this is the case, we might estimate substantial disparities between city-level effects of immigrants and natives absent any neighborhood effect heterogeneity, simply due to the uneven distribution of residents within cities.

To test this hypothesis, one could estimate neighborhood-level location effects and use these estimates to reconstruct the city-level effects as the equally weighted average of neighborhood effects.[24] This approach was taken in Card et al. (2022) to estimate industry-level wage

[22] Ideally, we would have estimated high school exposure effects instead of high school fixed effects. However, we resorted to estimating fixed effects due to limitations in available computing power.

[23] As a result, there are ten high schools in this grouped category.

[24] Unfortunately, we can't estimate the distribution of location effects at a neighborhood level due to our

premia as the average firm effects. We follow a similar approach. We estimate Equation (3) at a city level but reweight our regression inversely by the number of observations in each origin-destination(s) neighborhood cell, thereby equalizing the influence of each neighborhood on the aggregated city-level location effect.

Appendix Table A.3 presents the variance component estimates of natives and immigrants when estimating Equation (3) while reweighting by the number of observations in each origin-destination cell.[25] The first two rows display the standard deviation and Wald test of equality of location effects between natives and immigrants. We find two interesting results. First, columns 1 and 3 reveal that the standard deviation of the reweighted estimates is between 2 and 4 times larger than the unweighted estimates.[26] In Columns 2 and 4, we report the χ^2 test statistic and p-value for the test that all the location effects are equal and decisively reject the null.

Second, we find that even after accounting for the different spatial distribution of immigrants and natives, there is still substantial within-city heterogeneity. The correlation between the location effects of low-income immigrants and natives remains zero for low-income families and strongly positive for high-income families, aligned with our baseline estimates in Table 3. These results suggest that although the differential spatial distribution matters for the magnitude of city-level location effects, they do not explain the disparities between immigrants and natives.

Parental income of immigrants not reflecting earnings potential - Immigrants face earnings penalties due to frictions such as language barriers, cultural differences, and lack of network and information. Arellano-Bover and San (2023) estimate an immigrant-native earnings gap on arrival of 50%, which was fully closed only after 27-29 years. Therefore, if the heterogeneity in location effects with respect to parental income is driven by heterogeneity in skills, then parents' income rank is lower than their skill or ability rank would suggest, and we, therefore, classify high-earning potential parents as low earnings. By doing so, when we compare the location effects of immigrants and natives, we do not compare families with

data agreement restrictions.

[25]For more information on how we build the geographic units of neighborhoods, see Appendix H.

[26]This result of larger variance after accounting for sorting aligns with the finding is Card et al. (2022) in the context of wage equations, in which the reweighted firm premium was larger than the non-reweighted.

the same set of skills.[27]

To accommodate this, Appendix Table E.1 displays the correlation matrix of high and low-income families when estimating Equation 3, but instead of calculating parents' income rank in the national distribution, we do it within groups, therefore ranking parental among comparable groups. The standard deviation and correlations are qualitatively similar to those estimated in Table 2. This suggests that the negative relationship between low-income immigrants and natives is not driven mechanically by classifying immigrants with low income relative to natives but high income relative to other immigrants as low-income families.

Research design validation - In Appendix Section D, we provide several tests aimed at validating our research design and supporting our identification strategy. In Appendix Figures D.1 and D.2, we show a balancing exercise for both immigrants and natives of the relationship between the age at move and age at arrival to Israel and parents' years of schooling, as measured in the 1995 census. For native-born children, we also test for the relationship between parents' earnings growth when the children were young and the child's age at move. Our tests suggest that there is no statistically significant relationship between age at move and family characteristics conditional on the sequence of location choice fixed effects and the parents' income rank.

Linear location effects - The credibility of our approach depends additionally on the functional form assumptions in which location effects are linear with the years of exposure. In Appendix Section D, we provide several specification tests for that functional form assumption. Appendix Figure D.5 shows that, similar to the findings in the US (Chetty and Hendren, 2018a), Australia (Deutscher, 2020), and Canada (Laliberté, 2021), the relationship between years of exposure and the mean outcomes of children who spent all their childhood in the same location, are approximately linear, and find that in Israel A, the last year at which locations affect outcomes is approximately 18. This last finding aligns with the Israeli institutions in which most children enlist in the army immediately after high school. Additionally, in Appendix Figure D.8, we provide evidence that not only the relationship between mean outcomes and exposure time is linear but that the relationship between location effects themselves and exposure time is well approximated by a linear function.

[27] Although we do compare families with the same resources in childhood.

Robustness checks and sensitivity - We provide further robustness checks to assess the sensitivity of our estimates to several alternative specifications in Appendix section E. Appendix Table E.2 shows that our findings, and in particular the variance-covariance matrix of the location effects of immigrants and natives, are robust to alternative measures of income such as earnings and log earnings (excluding zeros). Appendix Section C compares the location effects estimated using only one-time movers, yielding similar qualitative results despite the smaller sample size, which allowed us to identify location effects for only 92 cities. Lastly, our current approach reweights cities based on the total number of families. Lastly, in Appendix Section E.3, we show the heterogeneity we report is robust to reweighting by the total number of movers (i.e., the total number of individuals who are included in our main regression sample) and to reweighting by group size.

8 The Distribution of Childhood Location Effects

Next, we extend our model and estimate the joint distribution of immigrant-native location effects. We use this extended model for decision-making for two diffrent tasks. First, we use the joint distribution to form the posterior mean effect of each city, which is the best location effect forecast that minimizes the mean squared error (James and Stein, 1961). Second, in Section 9, we exploit the joint distribution for housing policy exercise, in which we generate predictions of other features of the joint distribution. In the following section, we briefly discuss the additional assumptions we impose to estimate the joint distribution and the estimation procedure. For a more detailed discussion, see Appendix Section I.

8.1 Model and Estimation

We start with two main assumptions that are often employed in the literature. First, motivated by the central limit theorem, we assume that the estimates of location effects, $\hat{\theta}$, are conditionally normal. Second, as detailed in Appendix Section I, we find that each marginal distribution of θ_{jgp} is well approximated by a normal distribution by applying the log-spline estimator of Efron (2016). Therefore, we further assume that the prior distribution location effects is normally distributed.

To improve the predictive power of our model, we allow the mean location effects to vary linearly with a few of the city-level covariates z_j that were found to be predictive of location effects in Section 6.[28] Formally, we denote $z_j \in \mathbb{R}^p$ the vector of p covariates of city j (including a constant one for the intercept) where $z = (z_i, .., z_J)'$ is the corresponding $J \times p$ matrix. z_j includes the following covariates. First, following our findings from Section 6, it includes the city-level diversity index and locality welfare expenditure per capital. Since the diversity index is a function of the group shares of both immigrants and natives, it allows the location effects to correlate with the mobility patterns of both groups. Lastly, following the recent literature, our analysis additionally controls for log standard errors of $\hat{\theta}$ (Chen, 2022).

Therefore, for every income level p, the joint distribution of location effects $\theta_{jp} = (\theta_{j\mathcal{N}p}, \theta_{j\mathcal{I}p})$ takes the form:

$$\theta_{jp}|z_j, \Sigma \sim N(\mu_{\theta p}(z_j), \Omega_{\nu p})$$

where $\mu_{\theta p}(z_j) = (z_j'\beta_{\mathcal{N}p},\ z_j'\beta_{\mathcal{I}p})'$, and $\Omega_{\nu p}$ is the variance-covariance matrix of the signal with σ^2_{gp} the variance on diagonal and ρ_p the correlation coefficient. We estimate the model's hyperparameters in two steps. First, we estimate $\beta_p = (\beta_{\mathcal{N}p}, \beta_{\mathcal{I}p})$ by running a weighted least squares regression of $\hat{\theta}_{gp}$ on z separately for every group $g \in \{\mathcal{N}, \mathcal{I}\}$ and form each group's $J \times 1$ residual $r_{gp} = \hat{\theta}_{gp} - z\beta_{gp}$ and covariance matrix $\tilde{\Sigma} = M\Sigma M'$, where $M = I - z(z'z)^{-1}z'$ is the corresponding projection matrix. Then, similar to section 3.3, we estimate the city-size weighted unbiased variance component $\Omega_{\nu p}$ by method of moments, accounting for sampling error $\tilde{\Sigma}$.

Appendix Table A.4 reports the estimated hyperparameters of this extended model for families at the 25th and 75th percentile of the income distribution, with panels (i)-(iii) reporting the coefficients from β_{gp}, and panels (v)-(vi) reporting the variance components. As displayed in panels (i)-(iii), the coefficients on z_j closely match the findings from Section 6. Additionally, we find in panel (iv) a negative relationship between the location effects of native-born Israeli children and the log standard errors.

[28]Analogous to correlated random effect models, the ideal model would incorporate the relationship between the full origin-destination network and location effects. Due to limited access to computing power with the microdata, we take a simplified approach that allows only the mean to vary linearly with a few main components of the mobility network. While this is suboptimal, this approach can be rationalized by a decision that approximates the constrained oracle who only has access to the estimates and few features (Chen, 2023).

Our parsimonious extended model provides a good fit with high predictive power to each group's location effects. The first two rows of panel (vi), which report $\sqrt{\beta_g^2 \mathbb{V}(z) + \sigma_g^2}$, the total standard deviation of location effects, suggest that the variance of location effects of immigrants and natives qualitatively aligns with the Method of Moments estimates from Table 2. The last row in panel (vi) displays the total correlation between $\theta_{j\mathcal{N}p}$ and $\theta_{j\mathcal{I}p}$ which is the ratio of the total covariance between immigrants and natives ($\mathbb{C}ov(\theta_{j\mathcal{I}p}, \theta_{j\mathcal{N}p}) = \mathbb{C}ov(\mu_{\mathcal{N}}(z_j), \mu_{\mathcal{I}}(z_j)) + \rho\sigma_{\mathcal{N}}\sigma_{\mathcal{I}}$), divided by the product to the standard deviation reported in panel (vi). In line with our findings in Table 3, we find a small negative correlation between the location effects of low-income immigrants and natives and a strong positive correlation of 0.38 among high-income families. Lastly, to assess the predictive power of z_j, we calculate the R^2 of each group as the share of variation in θ explained by variation in z_j, $\frac{\beta_g^2 \mathbb{V}(z)}{\beta_g^2 \mathbb{V}(z) + \sigma_g^2}$, finding that z_j explains between 56-75% of the variation of native-born location effects and around 32-34% of the variation of immigrant location effects.

8.1.1 Posterior Mean

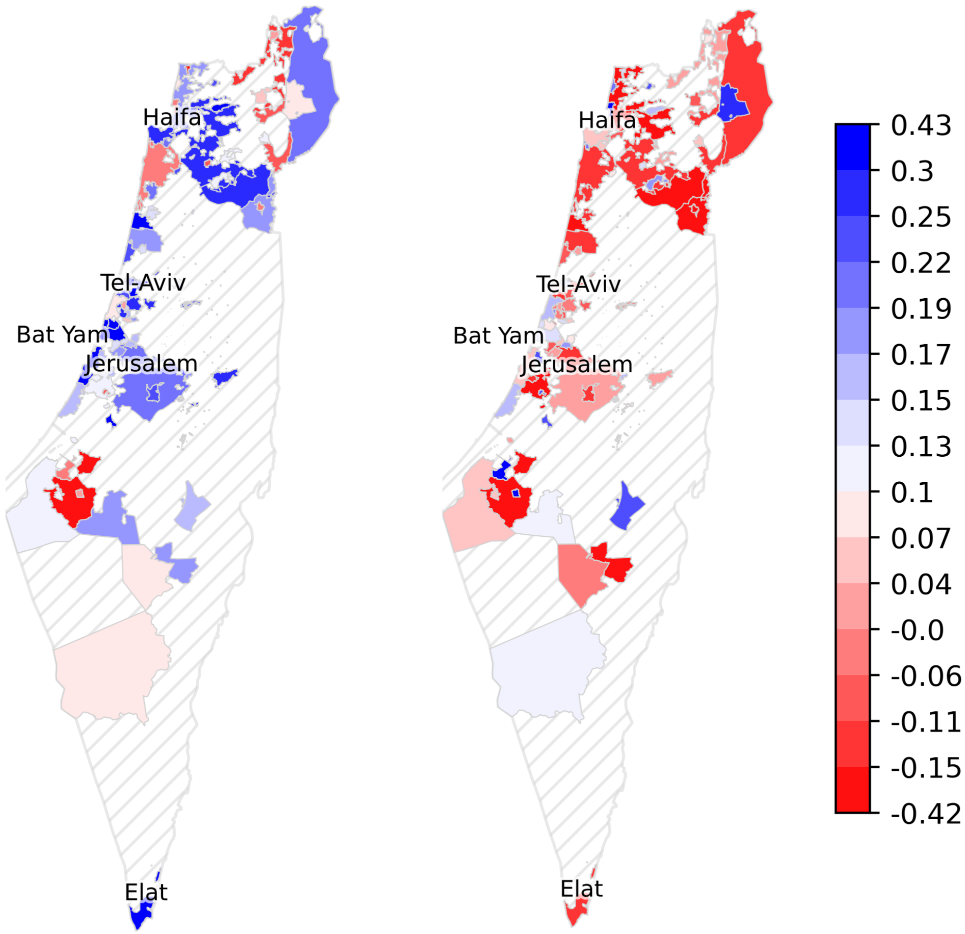
Using the estimated hyperparameters from columns (1) and (2) in Appendix Table A.4 as prior, we estimate the posterior mean,

$$\theta^*(z) \equiv \mathbb{E}[\theta|\hat{\theta}, z] = (\check{\Omega}_\nu^{-1} + \tilde{\Sigma}^{-1})^{-1} \left(\check{\Omega}_\nu^{-1} \check{\mu}_\theta(z) + \tilde{\Sigma}^{-1}\hat{\theta}\right), \tag{5}$$

which can be thought of as shrinking each of the estimated location effects $\hat{\theta}$ towards the linear prediction of $\hat{\theta}$ on z. Even if the true location effects are not normally distributed, the posterior mean yields a prediction of θ that reduces the mean square error at the cost of increased bias (James and Stein, 1961). Our approach aligns with other normal shrinkage methods used in the literature (Chetty et al., 2014a; Angrist et al., 2017; Chetty and Hendren, 2018b; Abdulkadiroğlu et al., 2020; Abaluck et al., 2021; Kline et al., 2022).

The location effects of low-income families - Figure 4 plots the posterior mean of the location effects across cities and regional councils in Israel for immigrants and natives from the 25th percentile of the national income distribution. Each effect describes the group's specific annual effect on income rank at age 28 compared to the annual effect of the average city. Table 4 displays the list of the top 10 and bottom 10 cities of low-income families, where in panel A, they are sorted by the immigrants' location effects, and in panel B, they

Figure 4: Posterior mean location effects, low-income families



(a) Natives, θ_{25} (b) Immigrants, θ_{25}

Note: These maps plot the children's posterior mean effect of year-long exposure to cities and regional councils in Israel on children's income rank at age 28 of low-income families whose parents are at the 25th percentile of the national income distribution. Figure (a) displays the effects for native-born children and Figure (b) displays the effects for immigrant children. The maps are constructed by grouping cities into 15 equally sized groups where the more blue the area, the greater its effect from the mean, and the more red the area, the smaller the effect compared to the mean.

are sorted by natives' location effects. In Appendix Table J.1, we provide the full lists of all the 98 cities and regional councils posterior means.[29]

The EB posterior means are highly variable both across cities and within cities across groups. The posterior mean effect of spending one more year in the worst city ranged between -0.42 lower yearly income rank at age 28 to 0.43 higher income rank, which are approximately a

[29]The corresponding full list of effects on high-income families is in Appendix Table J.2.

change of 642 ILS per year ($\approx$ 189 US dollars). Comparing immigrants and natives, it is apparent that there are significant differences between the cities that benefit one group and the cities that benefit the other group. Among low-income families, many of the northern Israeli cities are found to be places that benefit immigrants but not natives. The southern cities on the coastline of Israel, which have a high immigrant share, are also amongst the best cities for immigrants, while they are as good as the average for natives.

Table 4: Top and bottom cities based on the long-run location effects, low-income families

(A) Sorted by natives' loc. effects				(B) Sorted by immigrants' loc. effects			
Loc. name	Posterior mean imm. (1)	Posterior mean natives (2)	Share immigrants (3)	Loc. name	Posterior mean imm. (4)	Posterior mean native (5)	Share immigrants (6)
(i) Top 10							
Kokhav Ya'ir	-0.058	0.422	0.186	Netivot	0.431	-0.014	0.223
Hadera	-0.163	0.376	0.303	Ofaqim	0.417	0.007	0.299
Bat Yam	0.131	0.375	0.317	Akko	0.371	-0.044	0.277
Qiryat Motzkin	0.035	0.344	0.220	Qazrin	0.304	0.080	0.336
Qiryat Gat	0.226	0.340	0.327	Yavne	0.250	0.162	0.123
Elat	-0.136	0.336	0.180	Qiryat Gat	0.226	0.340	0.327
Ashdod	0.058	0.312	0.366	Qiryat Mal'akhi	0.224	-0.066	0.220
Rishon Leziyyon	0.147	0.310	0.187	Ma'alot-tarshiha	0.220	0.224	0.479
Karmi'el	0.169	0.307	0.393	Mateh Binyamin	0.220	0.174	0.115
Holon	0.097	0.304	0.177	Arad	0.219	0.165	0.419
Hod Hasharon	-0.084	0.304	0.071	Tirat Karmel	0.207	-0.052	0.197
	⋮	⋮	⋮		⋮	⋮	⋮
(ii) Bottom 10							
Shelomi	0.128	-0.422	0.202	Merhavim	-0.376	-0.268	0.000
Merhavim	-0.376	-0.268	0.000	Gedera	-0.299	0.158	0.116
Qiryat Arba	-0.069	-0.213	0.186	Emek HaMaayanot	-0.251	0.176	0.052
Upper Galilee	0.007	-0.141	0.134	Be'er Tuvia	-0.251	0.106	0.014
Migdal Haemeq	0.130	-0.079	0.289	Gilboa	-0.223	0.303	0.071
Qiryat Mal'akhi	0.224	-0.066	0.220	Bet She'an	-0.182	-0.013	0.074
Emek HaYarden	-0.136	-0.061	0.099	Dimona	-0.181	0.174	0.231
Tirat Karmel	0.207	-0.052	0.197	Mateh Asher	-0.165	0.174	0.000
Akko	0.371	-0.044	0.277	Hadera	-0.163	0.376	0.303
Hof HaCarmel	-0.147	-0.031	0.119	Misgav	-0.158	0.280	0.044

Note: This table presents the top 10 and bottom 10 locations in Israel with respect to the long-run effects on children's income rank at age 28 whose parents are from the 25th percentile of the income distribution. Panel (A) is sorted based on the posterior mean location effects of native-born children, and panel (B) is sorted based on the posterior mean location effects of immigrants. Columns 1-2 and 4-5 present the posterior mean location effect, and columns 3 and 6 present the share of former Soviet Union immigrants.

The location effects of high-income families - The posterior mean effects for families from the 75th percentile of the income distribution are plotted in Appendix Figure A.10 and reported in Appendix Table A.5. It is apparent that the lists of top and bottom cities for high-income native-born families share many of the localities with the immigrant high-income families' list.

9 Policy in the Face of Heterogeneity

The evidence on the importance of childhood residential location on children's long-term outcomes is the main motivation behind "moving to opportunity" policies, in which policymakers aim to motivate low-income housing voucher recipients to move to high-opportunity neighborhoods. Older literature emphasized selecting areas for public housing based on their poverty rates (Katz et al., 2001), while more recent studies suggest targeting locations based on children's outcomes in adulthood conditional on parental income (Bergman et al., 2019). We find that location effects in Israel exhibit substantial heterogeneity where the places that benefit low-income immigrants and native-born children are not necessarily the same places. Suppose we wanted to generate a list of recommended cities that provide the highest mobility for low-income children to inform housing policy in Israel, similar to Bergman et al. (2019). How does the treatment effect heterogeneity we documented affect the outcomes and design of the optimal policy? In this paper, we restrict attention to a model that maps closely to the selection of top places taken in the CMTO experiment. We focus on a partial equilibrium analysis and start with a simplified model that abstracts from capacity and budget constraints. In Section 9.6, we provide an extension that incorporates restrictions on possible behavioral responses.

9.1 Set-Up

Consider a decisionmaker whose task is to provide us with a single list of the top K cities in terms of their long-run effects on children's income in adulthood. Since public housing programs are targeted at low-income families, we restrict attention to a policy that takes into account only the long-run effects on low-income children. As such, to ease notation, we drop the p subscript that represents that heterogeneity.

We assume that the decisionmaker faces two main restrictions. First, due to ethical considerations, the decisionmaker is restricted to a decision rule that provides the same list to all groups, regardless of immigration status. This decision rule is described by the vector $\delta = (\delta_1, ..., \delta_J)'$ where $\delta_j \in \{0, 1\}$ indicates whether city j is selected by the policymaker. For example, in Bergman et al. (2019), the authors prespecified a list of neighborhoods that promote upward mobility, and then, in an experiment on housing voucher recipients, they recommended the families in the treatment group to move to one of the neighborhoods in their list.

Therefore, restricting the policy to be unified in this context implies precluding the possibility of providing different recommendations to different groups. Another example could be a decisionmaker who wants to choose K locations for new public housing installments for low-income families. Here, a unified decision rule implies that housing agencies cannot restrict access to an existing housing unit based on group characteristics.

Second, we assume that our decisionmaker faces uncertainty regarding the true value of θ and, therefore, can not form the decision δ based on the true location effects θ. While θ is unknown, we assume that, instead, the decisionmaker knows the distribution of θ and observes the estimates of θ and their variance, which we collect in the array $\mathcal{Y} = (\hat{\theta}, \Sigma)$. As a result, the decisionmaker forms decisions by minimizing the expected, rather true, loss, where the expectation is based on the posterior distribution of θ given the evidence $\mathcal{Y}$.

To evaluate the performance of different selection criteria and compare the gains of each group, we assume the decisionmaker evaluates the benefit of a decision rule relative to a first-best policy under full certainty about θ. The top K cities with the highest location effects of each group $g \in \{\mathcal{N}, \mathcal{I}\}$ are selected by the *oracle* rule:

$$\delta^*_{jgK} = \mathbb{1}\{\theta_{jg} \in \{\theta^{(1)}_{jg}, \theta^{(2)}_{jg}, ..., \theta^{(K)}_{jg}\}\}$$

where δ^*_{jgK} is the first best policy, and $\theta^{(l)}_{jg}$ is the l'th order statistic of the location effects of group g, i.e., the l'th largest value of θ_{jg}.

We define $\theta^*(\delta^*_{jgK}, K) \equiv \frac{1}{K}\sum_{j=1}^{J} \mathbb{E}[\theta_{jg}\delta^*_{jgK}]$ as the group g expected long-run effect of selected cities under the first best. Equipped with these definitions, the decisionmaker values the return of each city in comparison to the expected first-best value:

$$\vartheta_{jgK} = \theta^*(\delta^*_{jgK}, K) - \theta_{jg}. \tag{6}$$

Equation 6 gives the *regret* of not using the first best policy (Savage, 1954; Manski, 2004). It reflects the loss experienced by the decisionmaker once the true value of θ was known and if she didn't face any non-discrimination restriction.

9.2 Benchmark - Selection Based on the Average Effect

We start with a model rationalizing Bergman et al. (2019)'s selection criteria, where the goal of the decisionmaker is to choose the cities with the highest city-level location effects in the full population. Formally, the decisionmaker would like to choose the list of selected cities, δ, by minimizing the following loss function:

$$\mathcal{L}(\vartheta, \delta, \pi) = \sum_j \delta_j (\pi_{j\mathcal{I}} \vartheta_{j\mathcal{I}K} + (1 - \pi_{j\mathcal{I}}) \vartheta_{j\mathcal{N}K}) \tag{7}$$

subject to $\sum_{j=1}^{J} \delta_j = K$, where $\pi_{j\mathcal{I}} \in [0, 1]$ is the share of immigrants in city j in the data. This loss function implies that the decisionmaker would like to rank places based on the pooled city-level population mean effect, which describes how people sort within cities under the status quo:

$$\bar{\vartheta}_{jK} = \pi_{j\mathcal{I}} \vartheta_{j\mathcal{I}K} + (1 - \pi_{j\mathcal{I}}) \vartheta_{j\mathcal{N}K}, \tag{8}$$

and select the cities with the lowest $\bar{\vartheta}_{jK}$. Since immigrants are a minority group, the average city index assigns a small weight to their losses, disproportionally favoring the native-born group. With zero correlation between immigrants' and natives' location effects, by construction, places with low $\bar{\vartheta}_{jK}$ are more likely to be beneficial for natives but not necessarily beneficial to immigrants.

The decisionmaker does not observe the location effects θ_{jg} directly. Instead, we assume she treats the joint distribution from Section 8 as a prior and makes decisions based on the vector of location effects and their variance matrix $\mathcal{Y}$.[30] Therefore, the decisionmaker minimizes the expected loss, i.e., the Bayes risk, by choosing δ to minimize:

$$\begin{aligned} \mathcal{R}(\delta; \pi) &= \mathbb{E}[\mathcal{L}(\vartheta, \delta, \pi) | \mathcal{Y}] \\ &= \sum_j \delta_j (\pi_{j\mathcal{I}} \mathbb{E}[\vartheta_{j\mathcal{I}K} | \mathcal{Y}] + (1 - \pi_{j\mathcal{I}}) \mathbb{E}[\vartheta_{j\mathcal{N}K} | \mathcal{Y}]), \end{aligned}$$

where the expectation is taken over the posterior distribution of location effects given the evidence $\mathcal{Y}$, and $\mathbb{E}[\vartheta_{jgK} | \mathcal{Y}]$ is the posterior mean of the regret of city j and group $g \in \{\mathcal{N}, \mathcal{I}\}$.

[30] One might consider using the plug-in policy that replaces θ with $\hat{\theta}$. Manski (2021) refers to this approach as "as-if optimization" and establishes that it a dominated decision rules.

Therefore, the Bayesian decisionmaker ranks locations by their posterior expected regret:

$$\mathbb{E}[\bar{\vartheta}_{jK}|\mathcal{Y}] = \pi_{j\mathcal{I}}\mathbb{E}[\vartheta_{j\mathcal{I}K}|\mathcal{Y}] + (1 - \pi_{j\mathcal{I}})\mathbb{E}[\vartheta_{j\mathcal{N}K}|\mathcal{Y}],$$

and the optimal decision rule has the following form

$$\delta_{jK} = \mathbb{1}\{\mathbb{E}[\bar{\vartheta}_{jK}|\mathcal{Y}] \leq \kappa_K\},$$

with κ_K being the value of the Kth lowest posterior mean $\mathbb{E}[\bar{\vartheta}_{jK}|\mathcal{Y}]$. In what follows, we refer to this policy as targeting based on the *average status-quo* sorting patterns.

9.3 Accounting for Unknown Behavioral Response

We now turn to explore alternative policies that strive to avoid harming any of the groups that are being treated. Harm arises in our setting due two reasons. First, due to the inability of the decisionmaker to provide personalized recommendations ex-ante, together with the lack of information on which families will actually follow through and move to the recommended locations ex-post. Take-up uncertainty is a built-in restriction in the literature where suggested policies are based primarily on estimates of location effects but not on estimates of demand elasticities that incorporate information on recipients' compliance. This shortcoming was raised in Mogstad et al. (2023), who point out that there is no guarantee that families who received a recommendation in the CMTO experiment will sort into the places whose location effects are indeed high.[31] In this section, we propose a possible remedy by acknowledging the compliance uncertainty the decisionmaker is facing.

9.3.1 Who Shows Up?

Consider a scenario where the decisionmaker is uncertain about the identity of the families that move to each recommended location. Such a scenario could arise, for example, if the decisionmaker's task is to select a list of K cities for new public housing units, one unit in each chosen city. The loss function in this model mirrors Equation (7), where here, $\pi_{j\mathcal{I}}$ represents the probability that an immigrant family will eventually move into the housing unit built in city j. This probability is, therefore, a function of families' preferences, constraints,

[31] Similar concern was also raised in Pope and Sydnor (2011), who study statistical decision rules under anti-discrimination policies and note that the economic efficiency of such rules depends on individuals' behavioral responses.

information, and responses to the policy, all of which are unknown.

Facing this uncertainty, the decisionmaker can take several paths. Analogous to how the decisionmaker handles uncertainty with respect to each location effect θ, she can form a prior distribution on $\pi_{j\mathcal{I}}$ based on her beliefs. One justification for the decision rule in Equation (8) is that the decisionmaker's prior reflects a belief that public housing recipients sort according to the status quo, that is, similar to the existing sorting patterns within each city, regardless of the policy the face. We opt for a different approach, acknowledging our ignorance regarding family sorting behaviors. Our goal is to devise a policy that is robust to the least favorable compliance scenario: this is the *minimax* strategy (Wald, 1950), which was axiomatized by Gilboa and Schmeidler (1989).

Formally, given the vector of location effects θ, for every list of recommended cities δ, the least favorable compliance pattern implies that the worst-case regret is

$$\mathcal{L}^{\max(\mathcal{N},\mathcal{I})}(\theta,\delta) = \max_{\pi} \mathcal{L}(\vartheta,\delta,\pi) = \sum_j \delta_j \max_{\pi_j} \left\{\pi_{j\mathcal{I}}\vartheta_{j\mathcal{I}} + (1-\pi_{j\mathcal{I}})\vartheta_{j\mathcal{N}}\right\}. \tag{9}$$

If the decisionmaker knew the location effect of each city, they would minimize (9). However, with uncertainty regarding the true value of θ, the decisionmaker chooses δ to minimize the following expected maximum loss:

$$\mathcal{R}^{(\mathcal{N},\mathcal{I})}(\delta) = \mathbb{E}\left[\mathcal{L}^{\max(\mathcal{N},\mathcal{I})}(\theta,\delta)\middle|\mathcal{Y}\right] = \sum_j \delta_j \mathbb{E}\left[\max_{\pi_j}\left\{\pi_{j\mathcal{I}}\vartheta_{j\mathcal{I}} + (1-\pi_{j\mathcal{I}})\vartheta_{j\mathcal{N}}\right\}\middle|\mathcal{Y}\right] \tag{10}$$

subject to $\sum_{j=1}^{J}\delta_j = K$, where the expectation is taken over the posterior distribution of θ given the evidence $\mathcal{Y}$. Decisions motivated by optimizing objective functions involving both parameters that are not identified (π) and parameters that are point identified (θ) are sometimes referred to in the literature as robust Bayes decisions (Giacomini et al., 2021; Christensen et al., 2022). They are equivalent to a zero-sum game with nature, where nature knows the true location effects, and for every choice of recommended list of cities δ, it chooses the worst behavioral response π. By minimizing the maximum *regret*, the decisionmaker tries to achieve the oracle's first-best solution without violating the horizontal equity constraint.

Minimizing the objective function in Equation (10) yields the following decision rule in which the optimal policy is to rank locations based on their expected within-city posterior maximum regret:

$$\delta_{jK}^{(\mathcal{N},\mathcal{I})} = \mathbb{1}\{\mathbb{E}[\max\{\vartheta_{j\mathcal{I}K}, \vartheta_{j\mathcal{N}K}\}|\mathcal{Y}] \leq \kappa_K\}, \tag{11}$$

where κ_K is the maximum value such that there are exactly K cities with $\mathbb{E}[\max\{\vartheta_{j\mathcal{I}K}, \vartheta_{j\mathcal{N}K}\}|\mathcal{Y}] \leq \kappa_K$. Since this policy arises under uncertainty with respect to group identity, we refer to it as *minimax over $(\mathcal{N}, \mathcal{I})$*.

Connection to welfare economics - The spectrum of objectives between that implied by the observation-weighted loss function in Equation (7) and the minimax loss function in (9) map to the familiar social welfare criteria. At one extreme, Equation (7) can be thought of as a utilitarian social welfare function that linearly aggregates benefits across different groups, with each group's population share serving as the decisionmaker's social welfare weights. At the other extreme is the minimax decision loss function in Equation (9), which is equivalent to a Rawlsian decisionmaker with extreme equity preferences. The range of social preferences between the linear and the Rawlsian utility functions depends on the marginal rate of substitution between the two groups and reflects the decisionmaker's attitudes towards equity.

9.3.2 Who Shows Up and Where do They Go?

Next, we consider an example of a decisionmaker that faces uncertainty not only regarding the identity of each housing recipient but also regarding families' location choices. This model is inspired directly by the Bergman et al. (2019) experiment. In this framework, the decisionmaker recommends housing voucher recipients to relocate to one of the top K cities that provide high opportunities for low-income children. Then, given the recommended list δ, each family $g \in \{\mathcal{N}, \mathcal{I}\}$ sorts into cities according to the function $\pi_{jg}(\delta) \in [0, 1]$, such that $\sum_{j=1}^{J} (\pi_{j\mathcal{N}}(\delta) + \pi_{j\mathcal{I}}(\delta)) = 1$. Hence, the decisionmaker seeks to minimize the following loss function:

$$\begin{aligned}\mathcal{L}(\vartheta, \delta, \pi(\delta)) = \sum_{j=1}^{J} \Big[&\delta_j(\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}K} + \pi_{j\mathcal{N}}(\delta)\vartheta_{j\mathcal{N}K}) \\ &+ (1 - \delta_j)(\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}K} + \pi_{j\mathcal{N}}(\delta)\vartheta_{j\mathcal{N}K}) \Big]\end{aligned} \tag{12}$$

subject to $\sum_{j=1}^{K}\delta_j = K$. To avoid a degenerate minimax solution, we restrict attention to behavioral responses that satisfy full compliance, in which, given a selected list of recommended cities, recipients follow the recommendation and move to one of the cities in the list. We justify this approach following the findings of the CMTO experiment. First, the CMTO experiment suggests that Bergman et al. (2019) were able to build a technology that induces substantial compliance, increasing the share of families moving to recommended places by more than 38%. Second, Bergman et al. (2019) find that the sorting pattern of the control group in the CMTO experiment aligns with the sorting pattern in the status quo, absent the experiment. Therefore, the second part of the loss function in Equation (12) is likely constant, comprised of the share of non-compliers and the expected regret from the status quo sorting patterns.[32] Therefore, we proceed by restricting attention to the decisionmaker who would like to minimize the following loss function

$$\mathcal{L}(\vartheta,\delta,\pi(\delta)) = \sum_{j=1}^{J}\delta_j(\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}K} + \pi_{j\mathcal{N}}(\delta)\vartheta_{j\mathcal{N}K}) \tag{13}$$

subject to $\sum_{j=1}^{J}\delta_j = K$. We consider a decisionmaker who seeks a policy that is robust to the least favorable location choices. For given list of cities δ and location effects θ, the loss under such sorting behaviour is:

$$\mathcal{L}^{\max(\mathcal{N},\mathcal{I},city)}(\theta,\delta) = \max_{\pi(\delta)}\mathcal{L}(\vartheta,\delta,\pi(\delta)) = \max_{\pi(\delta)}\left\{\sum_{j=1}^{J}\delta_j(\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}K} + \pi_{j\mathcal{N}}(\delta)\vartheta_{j\mathcal{N}K})\right\},$$

where $\pi(\delta) = (\pi_{1\mathcal{N}}(\delta),\pi_{1\mathcal{I}}(\delta),...,\pi_{J\mathcal{N}}(\delta),\pi_{J\mathcal{I}}(\delta))'$. This loss arises when compliers belong to the immigration group with the highest regret and sort to the worst recommended city. Hence, the robust Baysean policy aims to select the δ that minimizes the expected loss given

[32] Formally, if we denote $\delta = 0$ as the status quo policy where no recommendation is introduced and represent the regret from the status quo policy as $\mathcal{L}(\vartheta,0) = \sum_{j=1}^{J}(\pi_{j\mathcal{I}0}\vartheta_{j\mathcal{I}K} + \pi_{j\mathcal{N}0}\vartheta_{j\mathcal{N}K})$, where π_{jg0} is the share of group g families in city j in the status quo, then the loss function of the decisionmaker is:

$$\mathcal{L}(\vartheta,\delta,\pi(\delta)) = \omega\sum_{j=1}^{J}\delta_j(\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}K} + \pi_{j\mathcal{N}}(\delta)\vartheta_{j\mathcal{N}K}) + (1-\omega)\mathcal{L}(\vartheta,0)$$

where ω is the share of compliers. In this model, the non-compliers will not affect the optimal policy of choosing K locations.

the evidence $\mathcal{Y}$:

$$\mathcal{R}^{(\mathcal{N},\mathcal{I},city)}(\delta) = \mathbb{E}[\mathcal{L}^{\max(\mathcal{N},\mathcal{I},city)}(\theta,\delta)|\mathcal{Y}]$$

subject to $\sum_j \delta_j = K$. This objective function yields a decision rule in which the optimal policy is to rank lists of locations of size K based on their expected maximum regret across all the cities in that list and across all groups:

$$\delta_K^{(\mathcal{N},\mathcal{I},\text{city})} = \arg\min_{\delta} \mathbb{E}[\max(\{\vartheta_{j\mathcal{N}K}, \vartheta_{j\mathcal{I}K}\}_{j\in S(\delta)})|\mathcal{Y}], \tag{14}$$

where $S(\delta) = \{j : \delta_j = 1\}$ is the set of recommended cities. Under this decision rule, the decisionmaker evaluates the posterior expectation of the maximum regret across all the selected locations and across immigrants and natives and chooses the list that attains the lowest worst-case regret. Therefore, hereafter, we refer to this policy as *minimax over ($\mathcal{N}$,$\mathcal{I}$,city)*.

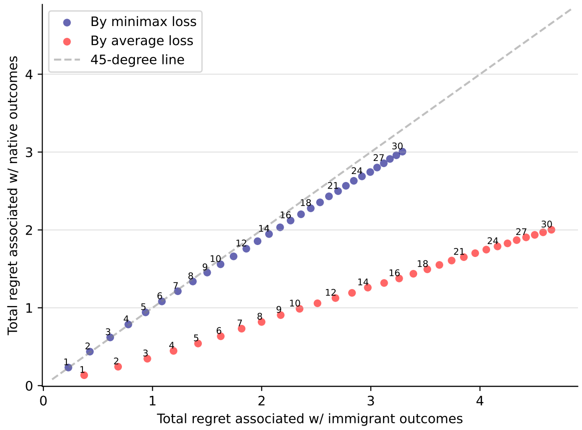
9.4 Evaluation of Each Policy

9.5 Tradeoffs

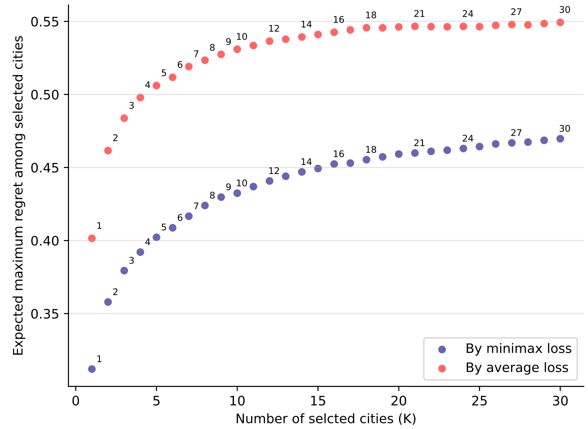
The tradeoff faced by the decisionmaker is visualized in Figure 5. We simulate the location effects for low-income families of the 98 cities implied by the mixing distribution in Appendix Table A.4. Using the simulation, we calculate the optimal first-best policy, and $\theta^*(\delta_g^*, K)$, and the average effect of the selected cities under the first-best policy, which is then used to compute the regret from not using the first-best, ϑ_{jgK}. Then, for every simulated draw and for a grid of values of K, we compute the mean status-quo average policy and the minimax policies. As detailed in Appendix K, we compute the empirical Bayes posterior expectation of the maximum by numerical integration. Lastly, for every policy, we compute the expected outcomes of each group by averaging across bootstrap draws.

Figure 5a plots the expected total regret of immigrants and natives from the status-quo average policy described in Equation (8) and the minimax ($\mathcal{N}/\mathcal{I}$) policy that ranks cities based on the within-city maximum regret described in Equation (11). Each dot corresponds to the sum of regret in selected cities by varying K between 1 and 30, and the corresponding K is displayed right next to the dot.

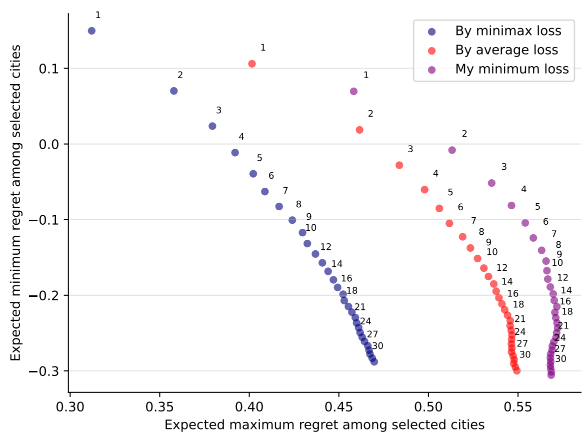
Figure 5: Regret under minimax and status-quo average targeting policies



(a) Total regret of immigrants and natives



(b) Max regret by number of selected cities



(c) Maximum and minimum regret

Note: These figures evaluate the regret losses of immigrants and natives from the average status-quo and minimax ($\mathcal{N}/\mathcal{I}$) policies. The regret, ϑ_{jgK}, of each city and group, immigrants and natives, is the difference between the location effect and the average location effects of the group-specific top K cities based on the true location effect θ. Blue dots plot the outcomes generated by the minimax ($\mathcal{N}/\mathcal{I}$) policy that ranks cities based on the within-city group maximum regret described in Equation (11). Red dots plot the outcomes generated by the status-quo average policy described in Equation (8), which ranks locations based on the within-city average regret, using the status quo group shares. In all the figures, the curves are generated by varying K, the number of selected cities, between 1 to 30, and the K is printed next to each dot. Figure 5a plots the total regret of immigrants and natives from policies that select the top K Israeli cities. Total regret is the sum of regret over selected cities, separately by immigration group, and the dashed line is the 45-degree line. Figure 5b plots the relationship between K, the number of selected cities, and the expected maximum regret. Figure 5c plots the expected maximum regret among selected cities against the expected minimum regret. The purple dots display the outcomes from an optimistic policy that rank places based on posterior expectation of the minimum regret between immigrants and natives $\mathbb{E}[\min(\{\vartheta_{j\mathcal{N}K}, \vartheta_{j\mathcal{I}K}\}_{j\in S(\delta)})|\mathcal{Y}]$.

For every K, a policy based on the status-quo average is clearly advantageous for natives while providing limited benefits to immigrants, which is reflected by the red curve lying below the 45-degree line. In contrast, for every K, the minimax strategy that ranks places based on the expected maximum regret lies very close to the 45-degree line, implying a more equitable outcome. We can see that as K increases, total regret increases, as it is harder to attain equal outcomes since there are fewer and fewer places that benefit both groups equally. While for a given K, the regret of native-born children under the minimax policy is higher than under the status-quo average, for every value of native-born child's regret, there exists a minimax strategy that attains the same level of regret, together with lower losses for immigrants.

Figure 5b plots the expected worst-case maximum regret ($\mathcal{I}/\mathcal{N}$) against K, the number of selected cities for the minimax ($\mathcal{I}/\mathcal{N}$) and average status quo policies. In line with the finding in Figure 5a, as K increases, the maximum worst-case regret increases. For every K, the average status quo policy results in a set of places that might end up with 20-90% higher losses compared to the minimax policy with the same K. As reflected in Equation (11), a choice of K corresponds to a choice of κ_K, the threshold of the posterior expectation of the maximum. Therefore, for a minimax decisionmaker, choosing K implies choosing the threshold for the worst possible regret, which reflects the decisionmaker's sensitivity to very bad outcomes. In this model, bad outcomes arise from two forms of uncertainty. First, similar to standard statistical decision problems, bad outcomes arise from estimation uncertainty. Additionally, bad outcomes arise here due to the heterogeneity in location effect. We further illustrate this point in the discussion of Table 5 below.

The ambiguity of the decisionmaker regarding compliance is reflected by their inability to determine the relative weights of each group's gains and how to aggregate these gains into a unified social objective. Under the worst-case scenario, which the minimax decision rule is trying to guard against, the families that comply with the treatment will sort into the least favorable places, maximizing the regret. In contrast, under the best-case scenario, it could also be possible that families who comply would be the ones that benefit the most, therefore further minimizing the regret of any given policy.

The relationship between the choice of K and the outcomes under the optimistic and pessimistic scenarios are presented in Figure 5c, which plots the expected maximum worst-case regret against the expected minimum best-case regret for the minimax (blue dots) and status-quo mean (red dots) policies for values of K between 1 to 30. To assess the performance of an optimistic decision maker when facing heterogeneity, we additionally analyze the policy that ranks places based on the posterior expectation of the minimum regret, $\mathbb{E}[\min(\{\vartheta_{j\mathcal{N}K}, \vartheta_{j\mathcal{I}K}\}_{j\in S(\delta)})|\mathcal{Y}]$, plotted by the purple dots. As noted by Hurwicz (1951), ranking places based on a convex average of the minimum and maximum loss, known as the α-minimax decision rule, reflects continuous types of decisionmakers with different levels of pessimism.

Figure 5c illustrates that in the face of heterogeneity, guarding against the least favorable scenarios could promise substantial gains also in states of the world that are less pessimistic. First, although for every K, the average status quo policy achieves lower values of minimum regret compared to the minimax policy, the worst- and best-case curves generated by the minimax decision rule outperform those of the average status quo decision rule. Interestingly, we find a similar pattern with the optimistic decision rule. By construction, for every K, the optimistic decision rule attains the lowest regret at the cost of risking with regret levels that are 2-4 times higher if families sort to the least favorable places. In contrast, the pessimistic minimax decision rule provides the lowest worst-case losses while maintaining close to zero (and lower) regret levels in the state of the world where families sort into the places that benefit them the most.[33]

9.5.1 Expected Benefits

Table 5 illustrates the costs and benefits of employing each of the three policies described in Section 9 for a given $K = 10$. All the values in the table are the Shekel value of the regret associated with an additional year of exposure. That is, they reflect the expected lost money from not residing in the expected group-specific first-best $\frac{1}{K}\sum_{j=1}^{J}\mathbb{E}[\vartheta_{jgK}\delta^*_{jgK}]$ city for one year. Columns 1-2 and 6-7 report the expected regret of immigrants and natives among selected cities; columns 3 and 8 report the expected within-city maximum regret $\frac{1}{K}\sum_{j=1}^{J}\max\{\vartheta_{j\mathcal{N}K}, \vartheta_{j\mathcal{I}K}\}\delta_j$ across immigrants and natives among selected cities; columns 4 and 9 report the expected maximum regret across all the cities, immigrants, and natives; and columns 5 and 10 report the share of selected cities that offer outcomes higher than the status-quo. In columns 1-4, we evaluate each of the policies mentioned above when the decisionmaker knows θ, and in columns 5-8, we consider a decisionmaker that doesn't know

[33] The optimistic decision rule attains negative regret levels because we defined the first best value as the expected average values across all the selected cities, while the optimistic scenario outperforms it and sends all the housing recipients to the single best place.

θ but relies on our estimated location effects and their standard errors $\mathcal{Y}$, as described in the previous section.

Table 5: Targeting trade-off from choosing the top 10 cities

	Known θ (oracle)					Unknown θ				
	Expected loss $\delta = 1$					Expected loss $\delta = 1$				
	Native born	Imm.	Max ($\mathcal{N}/\mathcal{I}$)	Max ($\mathcal{N}/\mathcal{I}/$ city)	Pr. both better status-quo	Native born	Imm.	Max ($\mathcal{N}/\mathcal{I}$)	Max ($\mathcal{N}/\mathcal{I}/$ city)	Pr. both better status-quo
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A) Low-inc. families										
i) Personalized policy										
Native-born	0.0	493.0	499.8	927.9	0.761	127.9	447.1	486.0	871.6	0.786
Immigrants	555.9	0.0	562.1	1088.2	0.501	480.5	128.6	513.3	936.9	0.591
ii) City level policy										
Average status quo	28.3	384.5	427.2	869.9	0.836	151.0	389.2	440.9	812.3	0.831
Minimax ($\mathcal{N}/\mathcal{I}$)	181.8	191.6	270.6	366.0	1.000	238.5	248.4	383.6	531.6	0.885
Minimax ($\mathcal{N}/\mathcal{I}$/city)	192.1	196.3	280.0	356.6	1.000	244.5	251.3	381.0	510.9	0.887
B) High-inc. families										
i) Personalized policy										
Native-born	0.0	450.7	474.7	1048.6	0.914	100.8	431.4	480.0	1035.9	0.917
Immigrants	288.2	0.0	309.5	691.5	0.903	258.6	145.4	339.6	676.0	0.935
ii) City level policy										
Average status quo	40.7	237.3	324.7	806.5	0.975	133.0	264.8	372.1	844.5	0.963
Minimax ($\mathcal{N}/\mathcal{I}$)	139.9	97.9	213.4	336.4	1.000	197.3	170.9	315.7	554.9	0.979
Minimax ($\mathcal{N}/\mathcal{I}$/city)	133.1	99.4	218.8	322.2	1.000	201.1	199.5	325.1	511.1	0.981

Note: This table evaluates the expected regret from different policies aiming at selecting the top 10 places in Israel. Regret is defined as the difference between the one-year location effects and the average benefits from the top cities of each group, immigrants and natives, in Shekel values (1 US \$ $\approx$ 3.4 ILS). It represents the lost earnings at age 28 from spending one year in the average selected city, compared to the average selected city under the first best policy. Columns 1-5 display the results from policies based on the true location effects, and columns 6-10 display the results from policies based on the expected values of location effects, where the expectation is taken over the posterior distribution of location effects conditional on the estimated location effects and standard errors, and the distribution of location effects from Table A.4 is treated as prior. Columns 1-2 and 6-7 report, for every group, immigrants and natives, the expected regret of each policy relative to the first-best policy. Columns 3 and 8 report the expected within-city immigrant-natives maximum regret among selected cities. Columns 4 and 9 present the expected maximum regret across all selected cities and immigrants and natives. Columns 5 and 10 present the share of selected places that provide benefits that are higher than the status quo for both immigrants and natives. The personalized policy is the policy that ranks locations based on the regret of each group. The table presents three city-level policies: The average status quo indicates the policy that ranks locations based on the city-level average regret as described in Equation (8). Minimax ($\mathcal{N}/\mathcal{I}$) indicates the policy that ranks locations based on the city-level immigrant-native maximum regret as described in Equation (11). Minimax ($\mathcal{N}/\mathcal{I}$/city) indicates the policy that ranks lists of 10 cities based on maximum regret across all the cities and groups in the list, as described in Equation (14).

As a result of the lack of correlation in location effects for immigrants and natives, policy recommendations based on effects for one group generate substantial regret for the other.

This can be seen in Panel (A.i), which presents the personalized first-best policy in which the decisionmaker recommends the top 10 locations based on the return of only one of the groups, either immigrants or natives. By construction, as shown in columns 1-2, under full information on location effects, the first best policy of each group generates no regret. Since location effects are not correlated, the average one-year regret for immigrants sent to the top locations for natives provides 555.9 ILS ($\approx$ \$163) lower income in adulthood compared to the first best. Similarly, sending native-born children to the top 10 immigrant places implies sending them to places that generate 493 fewer Shekels ($\approx$ \$145) in adulthood per year, compared to the first best.

When θ is unknown, the feasible policy is based on empirical Bayes shrinkage of functions of the estimated location effects, whereas in the personalized policy, the places are ranked according to the posterior mean of each location effect for every group. Columns 6-7 show that for immigrants (natives), the average recommended location under that feasible first-best generates 128 (127) ILS fewer Shekels in adulthood than the group's first-best. This loss illustrates the risk the decisionmaker faces from forming policy under uncertainty. If we provide the feasible first-best policy of one group to the other, their losses are as high as in columns 1 and 2.

Policies based on status quo mean effects place more weight on the gains of natives and, therefore, provide much weaker gains for immigrants. This is illustrated in Panel (A.ii), which shows that the status-quo average policy nearly attains the first-best outcomes of natives. Under full information about location effects, the status-quo average policy generates only 28 lower Shekels ($<$ \$10) for natives compared to the first best. Likewise, the Bayesian decisionmaker attains 151 lower Shekels per year ($\approx$\$45), which is only 18% more than the feasible first best. Accordingly, the returns of immigrants from the status quo average policy are almost as bad as the personalized recommendation that is based only on the gains of natives. They generate 78% of the losses for immigrants from the natives' oracle personalized recommendation and 82% of the losses under the Bayesian natives' personalized recommendation. Column 5 shows that even with full information on the location effects, under the status-quo average recommendation, almost 2 out of the 10 selected cities will end up with outcomes lower than the expected outcome under the status-quo sorting patterns for either of the groups.

While the status-quo average policy generates losses for the minority group, a minimax regret analysis shows it is possible to avoid extreme adverse outcomes for both groups. The second and third rows of panel (A.ii) display the average regret under the minimax $(\mathcal{N}, \mathcal{I})$ and $(\mathcal{N}, \mathcal{I}$, city) policies described in Equations (11) and (14) where the decisionmaker is ambiguous with respect to future behavioral responses. The second row reports the results for ranking places based on the maximum group regret, and the third row is based on the maximum regret across all the cities and groups. Under the feasible case where θ is unknown (columns 6-7), we can attain a more equal allocation that generates substantial improvements for immigrants. Comparing the minimax policies to the average status-quo policy, immigrants' regret can be reduced by almost half, at the cost of a 57% increase in the regret for natives. The expected maximum regret across the immigration groups (column 8) drops by 15%, and, remarkably, the expected maximum regret across groups and cities (column 9), i.e., the worst-case scenario, drops dramatically by 45%. Lastly, column 10 reports the share of selected cities that provide returns higher than the status quo for both groups. Due to the uncertainty regarding the values of the location effects, the minimax policies cannot provide us full insurance against cities that are worse than the status quo, as in column 5. Still, it ensures that at most 1.1 of the cities will not be beneficial to any groups.

The gains from employing the more equitable policies (i.e., the minimax policies) are less pronounced when the location effects of both groups are positively correlated. To illustrate this phenomenon, panel (B) reports the expected regret from such policies for high-income families who exhibit a strong correlation between the location effects of immigrants and natives. For immigrants, the city-level average posterior mean policy provides only 54% higher regret than the feasible first best reported in panel (A). This is much lower than the costs of low-income families, for which the city-level mean policy generates losses that are almost 3 times larger than the feasible first best. As a result, the minimax policies provide much lower improvements for immigrants than under the scenario of zero or negative correlation. The improvement in the share of places that are worse than the status quo is also marginal, and the average policy can ensure that more than 96% of the recommended cities will be better for both groups compared to the status quo.

9.5.2 Selected Israeli Cities

Table 6 displays the top 15 cities sorted according to their within-city posterior maximum regret across immigrants and natives (i.e., according to the minimax $(\mathcal{N},\mathcal{I})$ targeting policy) when $K = 10$. The regret is in Shekel values (1 US $\approx$ 3.4 ILS), and represents the lost earnings at age 28 from spending one year in city j, compared to the average city selected under the first best policy that allows for personalized recommendation and knows the true location effects.

The regret from selecting the top 10 leading cities is bounded for both groups, which implies that it is possible to identify at least a few places that provide substantial opportunities for both immigrants and natives. This is evident from columns 1-3, which report the posterior mean regret of each immigration group and the city average, weighted by the status quo within city shares. Compared to the personalized first best, an additional year spent in any of these cities generates, on average, 391-483 lower ILS earnings (115-143 US \$) for natives and lower 162-426 ILS earnings (47-126 US \$) for immigrants. Column (4) reports the posterior expectation of the maximum regret, ϑ_{jg5}, of each group $g \in \{\mathcal{N},\mathcal{I}\}$. Even in the least favorable scenario, children in the top five cities have incomes only 421-547 ILS lower (equivalent to approximately 123-160 US \$) than those in their average optimal city.

The top 10 cities based on the minimax $(\mathcal{N},\mathcal{I})$ targeting policy are likely to provide returns that are higher than those expected under the status quo sorting pattern. In column 5, we report the posterior probability that either the location effects of immigrants or of natives fall below the average location effect under the status quo sorting. This posterior probability is the analog of the false discovery rate (Benjamini and Hochberg, 1995) in multiple hypothesis testing settings where, for each city, we test the null that both the effects of immigrants and the effects of natives are greater than the expected value under the status quo sorting patterns. By averaging the first 10 values in column 10, we conclude that when selecting the top 10 cities following the minimax $(\mathcal{N},\mathcal{I})$ targeting policy, we should expect that at most 1 of these 10 cities would generate outcomes worse than the status quo for any of the two ethnic groups.

The last column indicates whether the city is selected under the model that allows for ambiguity with respect to both ethnicity and sorting patterns within the recommended list. When $K = 10$, the list of top 10 recommended cities coincides with the list of cities selected

Table 6: Top Israeli cities selected based on minimax criterion, $K = 10$

	Psterior mean					
	Native-born	Immigrants	Average	$\mathbb{E}[\max\{\vartheta_{\mathcal{N}}, \vartheta_{\mathcal{I}}\}\vert\mathcal{Y}]$	Worse than status-quo	Selected by minimax ($\mathcal{I}/\mathcal{N}$/city)
Loc. name	(1)	(2)	(3)	(4)	(5)	(6)
Qiryat Gat	391.7	162.1	237.2	421.7	0.039	Yes (749.4)
Ma'alot-tarshiha	399.6	339.5	368.3	499.7	0.103	Yes (749.4)
Karmi'el	477.8	212.5	316.7	503.1	0.108	Yes (749.4)
Rishon Leziyyon	512.3	208.6	265.4	517.0	0.066	Yes (749.4)
Yavne	354.2	435.0	425.0	526.8	0.118	Yes (749.4)
Mateh Binyamin	400.1	416.6	414.7	532.6	0.126	Yes (749.4)
Bat Yam	536.4	108.8	244.3	537.1	0.090	Yes (749.4)
Arad	401.2	430.0	417.9	542.8	0.142	Yes (749.4)
Ramla	451.4	439.9	443.0	544.7	0.106	Yes (749.4)
Ashqelon	483.0	426.5	446.3	546.9	0.086	Yes (749.4)
Ra'annana	537.2	321.0	350.4	575.5	0.213	
Qarne Shomeron	478.1	371.9	388.4	586.3	0.246	
Holon	587.9	217.1	282.7	590.2	0.187	
Be'er Sheva	574.9	407.0	458.4	596.8	0.144	
Tel-Aviv	487.1	564.4	554.6	600.9	0.142	

Note: This table reports a list of 15 Israeli cities sorted by the within-city posterior immigrant-native maximum regret. Regret is defined as the difference between the one-year location effects and the average benefits from the top cities of each group, immigrants and natives. It represents the lost earnings at age 28 from spending one year in city j, compared to the average city selected under the first best policy that allows for a personalized recommendation. Columns 1-3 report the posterior mean regret of native-born children, immigrants, and the average. Column 4 reports the posterior maximum regret across immigrants and natives. Column 5 reports the posterior probability that the location effects of immigrants or the location effects of native-born are lower than the average effect under the status quo sorting patterns. Column 6 reports which cities are selected as the top 10 cities based on the minimax (N/I/city) policy that ranks lists of 10 cities based on their posterior maximum regret across all cities and groups, where the posterior maximum regret of the selected list is presented in parentheses.

using the minimax $(\mathcal{N}, \mathcal{I})$ targeting policy in which ambiguity regarding behavioral responses was ignored. In parenthesis, we present the posterior expectation of the maximum regret across all the selected cities and across both immigrants and natives ϑ_{jg105}. Consistent with Jensen's inequality, the expected maximum value across all cities exceeds the average values of the top 10 cities found in column 4.

Robustness to location effect normalization - The regret normalization provides us with a measure that compares each policy to the non-restricted optimal one. Additionally, it enables us to overcome the consequences of our estimation procedure in which only relative effects, rather than exact levels, are identified.[34] To assess the sensitivity of the results to the regret normalization, Appendix Table A.6 replicates Table 6 while normalizing the value of

[34] A decisionmaker who seeks to minimize regret is concerned not only about the outcome she receives but also about the outcome she would have received had she chosen differently. One of our motivations for using the regret normalization is a discussion we had with Israeli government officials who expressed their disappointment after finding the excess heterogeneity in location effects in Israel and raised the concern that the first-best personalized policy will never be feasible.

each place in comparison to that expected under the status quo sorting patterns. Selecting the top 10 Israeli cities using the mean status quo normalization ends up with the same list as in Table 6, although the within-list ranking is different.

9.6 Model Extension

The stylized models depicted in Section 9.3 provide a clear, easy-to-interpret closed-form decision rule for the minimax decisionmaker who seeks to be robust against the least favorable sorting scenario. Nevertheless, these models were derived under simplified assumptions that might not hold in reality. First, the minimax decisionmaker behaves as if all the families might sort into a single worst place – a phenomenon that is rejected by the data. To illustrate that, Appendix Figure A.11 plots the distribution of location choices of families if they follow the minimax $\mathcal{N}/\mathcal{I}/city$ strategy and face the minimax $\mathcal{N}/\mathcal{I}/city$ decisionmaker. This figure shows that the minimax behavior assumes families sort into a small set of places, which might not be a reasonable belief. Second, the models in Section 9.3 do not take into account capacity constraints and other restrictions that might arise in real-world problems. Therefore, in this section, we describe our extended model that restricts the sorting probabilities to align better with the data while maintaining compliance uncertainty.

9.6.1 Restricting the Model for Location Choices

Let $\pi_{jg0} \in (0,1)$ be the status quo share of group $g \in \{\mathcal{N}, \mathcal{I}\}$ individuals who live in city j in the absent of any policy such that $\sum_j (\pi_{j\mathcal{N}0} + \pi_{j\mathcal{I}0}) = 1$. We impose the following restrictions on the location choice probabilities of natives, $\pi_{\mathcal{N}}(\delta) = (\pi_{j\mathcal{N}}(\delta), ..., \pi_{J\mathcal{N}}(\delta))'$, and immigrants $\pi_{\mathcal{I}}(\delta) = (\pi_{j\mathcal{I}}(\delta), ..., \pi_{J\mathcal{I}}(\delta))'$. First, we rule out sorting probabilities that deviate substantially from the status-quo sorting patterns by considering choice probabilities $\pi(\delta) = (\pi_{\mathcal{N}}(\delta)', \pi_{\mathcal{I}}(\delta)')'$ whose distance from the status quo sorting $\pi_0 = (\pi'_{0\mathcal{N}}, \pi'_{0\mathcal{I}})'$ is bounded. To measure the distance between $\pi(\delta)$ and π_0, we use the Total Variation distance function, which gives the largest absolute difference between the probability distributions across all the cities:

$$TV_{\pi_0}(\pi(\delta)) = \sup_{(j,g) \in \{1,..,J\} \times \{\mathcal{N}, \mathcal{I}\}} |\pi_{jg}(\delta) - \pi_{0jg}|.$$

With this metric, $TV_{\pi_0}(\pi) = 0$ implies that all the families comply according to the status quo distribution, while as $TV(\pi) \to 1$, all the families belong to a single immigration group and sort into a single place. Our aim is to examine the optimal recommendation policy under a hypothetical bound on the tendency of families to deviate from the status-quo shares:

$$TV(\pi) \leq a,$$

where $a \in [0, 1]$ represents the degree of unexpected sorting behavior. This restriction, together with the logical bound of $\pi_{jg}(\delta) \in [0, 1]$, and the restriction of $\pi_{jg}(\delta) = 0$ if $\delta_j = 0$ we describe in Section 9.3, implies that location choices are set identified, satisfying for every city j with $\delta_j = 1$ and for every $g \in \{\mathcal{N}, \mathcal{I}\}$:

$$\pi_{jg}(\delta) \in [\max\{\tilde{\pi}^{\delta}_{jg0} - a, 0\}, \min\{\tilde{\pi}^{\delta}_{jg0} + a, 1\}], \quad \text{with} \quad \tilde{\pi}^{\delta}_{jg0} = \frac{\pi_{jg0}}{\sum_j (\pi_{j\mathcal{N}0} + \pi_{j\mathcal{I}0})\delta_j}, \tag{15}$$

where $\tilde{\pi}^{\delta}_{jg0}$ is the status quo shares normalized to sum to one across selected cities. These restrictions ensure that location choices follow approximately the status quo distribution while maintaining ambiguity regarding compliance ("where do they go?") and families' group affiliation ("who shows up?") governed by the parameter $a > 0$.

With uncertainty regarding location choices and families' identity, the minimax decisionmaker would like to choose δ that is robust to the least favorable behavioral responses. For any given δ, the maximum regret can be written as:

$$\begin{aligned} \mathcal{L}^{max}_R(\vartheta, \delta) = &\max_{\pi(\delta)} \mathcal{L}(\vartheta, \delta, \pi(\delta)) \\ &\text{s.t} \quad \text{Equations (15)}, \\ &\sum_j (\pi_{j\mathcal{N}}(\delta) + \pi_{j\mathcal{I}}(\delta)) = 1, \end{aligned} \tag{16}$$

where $\mathcal{L}(\vartheta, \delta, \pi(\delta))$ is defined in Equation (13) and the decisionmaker, therefore, chooses the δ that minimizes:

$$\mathcal{R}^{\mathcal{N},\mathcal{I},city}_R(\delta) = \mathbb{E}[\mathcal{L}^{max}_R(\vartheta, \delta)|\mathcal{Y}] \tag{17}$$

subject to $\sum_j \delta_j = K$. Similar to the policy in Equation (14), the decision rule in Equation

(17) ranks lists of size K places based on the expected maximum regret under the least favorable compliance. Unlike the unrestricted model, here, the decisionmaker assumes that there is a distribution of families across all recommended cities, ruling out the possibility that all the families sort to a single least beneficial location. The smaller the value of a, the more location choices align with the status-quo sorting pattern. When $a = 0$, lists of places are ranked based on the posterior average status quo regret $\sum_j \delta_j \mathbb{E}[\tilde{\pi}^{\delta}_{j\mathcal{N}0}\vartheta_{j\mathcal{N}10} + \tilde{\pi}^{\delta}_{j\mathcal{I}0}\vartheta_{j\mathcal{I}10}]$. In contrast, when $a \to 1$, the decisionmaker faces more uncertainty regarding families behavioral responses, and the optimal decision approaches the one reported in Equation (14).

In Appendix Section L, we show that these location choices can be microfunded by a discrete choice model with additively separable components of compliance and private evaluation, and where their private valuations are not restricted to follow a particular distribution or correlation structure across cities. This model implies a rational behavior accompanied by flexible heterogeneity in compliance responses.

Estimation - Following Christensen et al. (2022) and as detailed in Appendix Section K, we estimate the decision rule implied by Equation (17) using a bootstrap implementation. Given a value of $a \in [0, 1]$, we estimate the bootstrap average maximum risk:

$$\mathcal{R}^{*\mathcal{N},\mathcal{I},city}_{R}(\delta) = \mathbb{E}^{*}[\mathcal{L}^{max}(\vartheta, \delta)|\mathcal{Y}], \tag{18}$$

where the expectation operator $\mathbb{E}^{*}$ denotes the expectation with respect to the S bootstrap draws from the posterior distribution of θ given $\mathcal{Y}$, and in each bootstrap simulation, we solve Equation (16) by linear programing. The minimax bootstrap decision rule is then the policy that attains the minimal expected maximum risk:

$$\delta^{*\mathcal{N},\mathcal{I},city}_{K,R} = \arg\min_{\delta} \mathcal{R}^{*\mathcal{N},\mathcal{I},city}_{R}(\delta).$$

Under strictly positive prior, Christensen et al. (2022) establish the optimality of the bootstrap decision rule. To avoid computing $\binom{J}{K}$ values of the expected maximum regret in Equation (18), we reduce the dimensionality of our simulation by taking a recursive approach. See Appendix Section K for further details.

Table 7 reports the top 10 selected cities and the posterior mean regret for both immigrants and natives for values of $a = 0.005$ and $a = 0.9$. For $a = 0.005$, with location choice restricted to align closely with the status quo distribution, the selected cities provide significant benefits (low regret levels) for native-born families, as they constitute the majority in each city according to the status quo. In contrast, for $a = 0.9$, with fewer restrictions on location choices, the selected cities offer more equal outcomes for both groups. The average regret for each city does not exceed 483 lost shekels per year compared to the oracle first-best policy.

Table 7: Top 10 Israeli cities selected based on the restricted minimax criterion

$\alpha = 0.001$			$\alpha = 0.9$		
Loc. name	Post. mean imm. (1)	Post. mean natives (2)	Loc. name	Post. mean imm. (3)	Post. mean native (4)
Kokhav Ya'ir	826.1	37.5	Qiryat Gat	391.7	162.1
Qiryat Gat	391.7	162.1	Ma'alot-tarshiha	399.6	339.5
Bat Yam	536.4	108.8	Karmi'el	477.8	212.5
Hod Hasharon	865.8	217.2	Rishon Leziyyon	512.3	208.6
Rishon Leziyyon	512.3	208.6	Yavne	354.2	435.0
Qiryat Motzkin	683.9	156.7	Mateh Binyamin	400.1	416.6
Jezreel Valley	951.2	233.4	Bat Yam	536.4	108.8
Gilboa	1078.4	218.6	Arad	401.2	430.0
Holon	587.9	217.1	Ramla	451.4	439.9
Misgav	978.0	254.3	Ashqelon	483.0	426.5

Note: This table reports the list of 10 selected cities by the restricted minimax ($\mathcal{N}/\mathcal{I}/city$) decision maker depicted in Section 9.6. Columns 1-2 report the posterior mean regret of the selected 10 cities when $a = 0.01$, which reflects a prior with little deviation from the status-quo sorting patterns, and columns 3-4 report the posterior mean regret of the selected 10 cities when $a = 0.9$, which reflects very very little restrictions on sorting patterns. Regret is defined as the difference between the one-year location effects and the average benefits from the top cities of each group, immigrants and natives. It represents the lost earnings at age 28 from spending one year in city j, compared to the average city selected under the first best policy that allows for a personalized recommendation.

10 Conclusion

This paper studies the heterogeneity in the causal effect of cities and regional councils in Israel on children's income in adulthood. We exploit the unexpected mass migration wave from the former Soviet Union in the 90s and estimate each city's effect separately for immigrants and natives. Our exploration into the nuanced association between childhood location effects of natives vs. immigrants in Israeli cities has illuminated the surprising fact that cities that benefit one group are not necessarily the cities that benefit the other groups.

In light of these findings, housing policies that recommend locations based on population-wide means can disproportionately harm minorities. We discuss the trade-offs policymakers face when implementing a unified policy that cannot be conditioned on individual characteristics when uncertainty with respect to the true effects and with respect to future compliance patterns are in place. Using the decision-theoretic framework, we show that by acknowledging the ambiguity with respect to individuals' sorting, it is possible to find at least 5 cities in Israel that are beneficial to both groups.

In this paper, we examine the tradeoff that a decision-maker faces when a policymaker cannot discriminate based on certain individual characteristics. While previous literature has emphasized the optimality of using all predetermined characteristics to achieve fairness and efficiency (Kleinberg et al., 2018b; Cowgill and Tucker, 2019; Rambachan et al., 2020; Liang et al., 2021), such policies are not always feasible due to non-discriminatory laws or inherent restrictions (Chan and Eyster, 2003; Ellison and Pathak, 2021). Our model demonstrates that we can improve the fairness of the restricted policy by modeling the uncertainty generated by such restrictions using a decision-theoretic framework. Therefore, our model is relevant to various other settings where the treatment is not directed to individuals but to predefined groups. Such domains include teacher and school assignment (Biasi et al., 2021; Abdulkadiroğlu et al., 2020; Rose et al., 2022; Bates et al., 2022), admission policies (Ellison and Pathak, 2021), job training programs (Card et al., 2018), or criminal justice (Kleinberg et al., 2018a; Agan and Starr, 2018), and health care decisions (Obermeyer et al., 2019).

References

(1998). *Local Authorities in Israel.* Publication No. 1134. Central Bureau of Statistics.

Abaluck, J., Caceres Bravo, M., Hull, P., and Starc, A. (2021). Mortality effects and choice across private health insurance plans. *The quarterly journal of economics*, 136(3):1557–1610.

Abdulkadiroğlu, A., Pathak, P. A., Schellenberg, J., and Walters, C. R. (2020). Do parents value school effectiveness? *American Economic Review*, 110(5):1502–39.

Abramitzky, R., Baseler, T., and Sin, I. (2022). Persecution and migrant self-selection: Evidence from the collapse of the communist bloc. Technical report, National Bureau of Economic Research.

Abramitzky, R., Boustan, L. P., and Connor, D. (2020). Leaving the enclave: Historical evidence on immigrant mobility from the industrial removal office. Technical report, National Bureau of Economic Research.

Abramitzky, R., Boustan, L. P., Jácome, E., and Pérez, S. (2019). Intergenerational mobility of immigrants in the us over two centuries. Technical report, National Bureau of Economic Research.

Agan, A. and Starr, S. (2018). Ban the box, criminal records, and racial discrimination: A field experiment. *The Quarterly Journal of Economics*, 133(1):191–235.

Alexander, R. and Ward, Z. (2018). Age at arrival and assimilation during the age of mass migration. *The Journal of Economic History*, 78(3):904–937.

Aliprantis, D., Martin, H., and Tauber, K. (2024). What determines the success of housing mobility programs? *Journal of Housing Economics*, page 102009.

Altonji, J. G. and Card, D. (2018). The effects of immigration on the labor market outcomes of less-skilled natives. In *The New Immigrant in the American Economy*, pages 137–170. Routledge.

Andrews, I., Bowen, D., Kitagawa, T., and McCloskey, A. (2022). Inference for losers.

Angrist, J. and Lavy, V. (2009). The effects of high stakes high school achievement awards: Evidence from a randomized trial. *American economic review*, 99(4):1384–1414.

Angrist, J. D., Hull, P. D., Pathak, P. A., and Walters, C. R. (2017). Leveraging lotteries for school value-added: Testing and estimation. *The Quarterly Journal of Economics*, 132(2):871–919.

Arellano-Bover, J. and San, S. (2023). The role of firms and job mobility in the assimilation of immigrants: Former soviet union jews in israel 1990-2019.

Ba, B., Bayer, P., Rim, N., Rivera, R., and Sidibé, M. (2021). Police officer assignment and neighborhood crime. Technical report, National Bureau of Economic Research.

Bai, Y., Santos, A., and Shaikh, A. M. (2022). A two-step method for testing many moment inequalities. *Journal of Business & Economic Statistics*, 40(3):1070–1080.

Bartel, A. P. (1989). Where do the new us immigrants live? *Journal of Labor Economics*, 7(4):371–391.

Bates, M. D., Dinerstein, M., Johnston, A. C., and Sorkin, I. (2022). Teacher labor market equilibrium and the distribution of student achievement. Technical report, National Bureau of Economic Research.

Beaman, L. A. (2012). Social networks and the dynamics of labour market outcomes: Evidence from refugees resettled in the us. *The Review of Economic Studies*, 79(1):128–161.

Benjamini, Y. and Hochberg, Y. (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, P., Chetty, R., DeLuca, S., Hendren, N., Katz, L. F., and Palmer, C. (2019). Creating moves to opportunity: Experimental evidence on barriers to neighborhood choice. Technical report, National Bureau of Economic Research.

Bertrand, M., Luttmer, E. F., and Mullainathan, S. (2000). Network effects and welfare cultures. *The quarterly journal of economics*, 115(3):1019–1055.

Biasi, B., Fu, C., and Stromme, J. (2021). Equilibrium in the market for public school teachers: District wage strategies and teacher comparative advantage. Technical report, National Bureau of Economic Research.

Bobba, M., Ederer, T., Leon-Ciliotta, G., Neilson, C., and Nieddu, M. G. (2021). Teacher compensation and structural inequality: Evidence from centralized teacher school choice in perú. Technical report, National Bureau of Economic Research.

Buchinsky, M., Gotlibovski, C., and Lifshitz, O. (2014). Residential location, work location, and labor market outcomes of immigrants in israel. *Econometrica*, 82(3):995–1054.

Card, D. (2009). Immigration and inequality. *American Economic Review*, 99(2):1–21.

Card, D., Kluve, J., and Weber, A. (2018). What works? a meta analysis of recent active labor market program evaluations. *Journal of the European Economic Association*, 16(3):894–931.

Card, D., Rothstein, J., and Yi, M. (2022). Industry wage differentials: A firm-based approach. *Unpublished draft, University of California, Berkeley.*

Chan, J. and Eyster, E. (2003). Does banning affirmative action lower college student quality? *American Economic Review*, 93(3):858–872.

Chen, J. (2022). Empirical bayes when estimation precision predicts parameters. *arXiv preprint arXiv:2212.14444*.

Chen, J. (2023). Empirical bayes when estimation precision predicts parameters.

Chetty, R., Friedman, J. N., Hendren, N., Jones, M. R., and Porter, S. R. (2018). The opportunity atlas: Mapping the childhood roots of social mobility. Technical report, National Bureau of Economic Research.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014a). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American economic review*, 104(9):2593–2632.

Chetty, R. and Hendren, N. (2018a). The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects. *The Quarterly Journal of Economics*, 133(3):1107–1162.

Chetty, R. and Hendren, N. (2018b). The impacts of neighborhoods on intergenerational mobility II: County-level estimates. *The Quarterly Journal of Economics*, 133(3):1163–1228.

Chetty, R., Hendren, N., and Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, 106(4):855–902.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014b). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics*, 129(4):1553–1623.

Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big data*, 5(2):153–163.

Christensen, T., Moon, H. R., and Schorfheide, F. (2022). Optimal discrete decisions when payoffs are partially identified. *arXiv preprint arXiv:2204.11748*.

Chyn, E., Collinson, R., and Sandler, D. (2022a). The long-run effects of residential racial desegregation programs: Evidence from gautreaux.". Technical report, Working Paper.

Chyn, E., Haggag, K., and Stuart, B. A. (2022b). The effects of racial segregation on intergenerational mobility: Evidence from historical railroad placement. Technical report, National Bureau of Economic Research.

Chyn, E. and Katz, L. F. (2021). Neighborhoods matter: Assessing the evidence for place effects. *Journal of Economic Perspectives*, 35(4):197–222.

Cohen-Goldner, S. and Paserman, M. D. (2011). The dynamic impact of immigration on natives' labor market outcomes: Evidence from israel. *European Economic Review*, 55(8):1027–1045.

Cowgill, B. and Tucker, C. E. (2019). Economics, fairness and algorithmic bias. *preparation for: Journal of Economic Perspectives*.

Cutler, D. M. and Glaeser, E. L. (1997). Are ghettos good or bad? *Quarterly Journal of Economics*, 112:827–872.

Damm, A. P. (2009). Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence. *Journal of Labor Economics*, 27(2):281–314.

Damm, A. P. and Dustmann, C. (2014). Does growing up in a high crime neighborhood affect youth criminal behavior? *American Economic Review*, 104(6):1806–32.

Deutscher, N. (2020). Place, peers, and the teenage years: long-run neighborhood effects in australia. *American Economic Journal: Applied Economics*, 12(2):220–49.

Dustmann, C., Schönberg, U., and Stuhler, J. (2016). The impact of immigration: Why do studies reach such different results? *Journal of Economic Perspectives*, 30(4):31–56.

Edin, P.-A., Fredriksson, P., and Åslund, O. (2003). Ethnic enclaves and the economic success of immigrants—evidence from a natural experiment. *The quarterly journal of economics*, 118(1):329–357.

Efron, B. (2016). Empirical bayes deconvolution estimates. *Biometrika*, 103(1):1–20.

Ellison, G. and Pathak, P. A. (2021). The efficiency of race-neutral alternatives to race-based affirmative action: Evidence from chicago's exam schools. *American Economic Review*, 111(3):943–75.

Eshaghnia, S. (2023). Is zip code destiny? Technical report.

Fogel, N. (2006). Crime in israel – 2002, potential expansion of the cbs crime statistics. Technical Report 20, Central Bureau of Statistics – Social-Economic Research Department.

Garin, A. and Rothbaum, J. (2022). The long-run impacts of public industrial investment on regional development and economic mobility: Evidence from world war ii. Technical report, Working Paper.

Giacomini, R., Kitagawa, T., and Read, M. (2021). Robust bayesian analysis for econometrics.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18:141–153.

Goldner, S. C., Eckstein, Z., and Weiss, Y. (2012). *Immigration and Labor Market Mobility in Israel, 1990 to 2009*. Mit Press.

Gould, E. D., Lavy, V., and Paserman, M. D. (2004). Immigrating to opportunity: Estimating the effect of school quality using a natural experiment on ethiopians in israel. *The Quarterly Journal of Economics*, 119(2):489–526.

Gould, E. D., Lavy, V., and Paserman, M. D. (2011). Sixty years after the magic carpet ride: The long-run effect of the early childhood environment on social and economic outcomes. *The Review of Economic Studies*, 78(3):938–973.

Graham, B. S., Ridder, G., Thiemann, P., and Zamarro, G. (2023). Teacher-to-classroom assignment and student achievement. *Journal of Business & Economic Statistics*, 41(4):1328–1340.

Gu, J. and Koenker, R. (2020). Invidious comparisons: Ranking and selection as compound decisions. *arXiv preprint arXiv:2012.12550*.

Heckman, J. and Landersø, R. (2021). Lessons for americans from denmark about inequality and social mobility. *Labour Economics*, page 101999.

Hurwicz, L. (1951). The generalized bayes minimax principle: a criterion for decision making under uncertainty. *Cowles Comm. Discuss. Paper Stat*, 335:1950.

James, W. and Stein, C. M. (1961). Estimation with quadratic loss. 1:361–380.

Katz, L. F., Kling, J. R., and Liebman, J. B. (2001). Moving to opportunity in Boston: Early results of a randomized mobility experiment. *The Quarterly Journal of Economics*, 116(2):607–654.

Kitagawa, T. and Tetenov, A. (2018). Who should be treated? empirical welfare maximization methods for treatment choice. *Econometrica*, 86(2):591–616.

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., and Mullainathan, S. (2018a). Human decisions and machine predictions. *The quarterly journal of economics*, 133(1):237–293.

Kleinberg, J., Ludwig, J., Mullainathan, S., and Rambachan, A. (2018b). Algorithmic fairness. In *Aea papers and proceedings*, volume 108, pages 22–27. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Kline, P., Rose, E. K., and Walters, C. R. (2022). Systemic discrimination among large us employers. *The Quarterly Journal of Economics*, 137(4):1963–2036.

Kline, P., Rose, E. K., and Walters, C. R. (2023). A discrimination report card. *arXiv preprint arXiv:2306.13005*.

Kline, P. and Walters, C. (2021). Reasonable doubt: Experimental detection of job-level employment discrimination. *Econometrica*, 89(2):765–792.

Kline, P. M., Rose, E. K., and Walters, C. R. (2024). A discrimination report card. Technical report, National Bureau of Economic Research.

Laliberté, J.-W. (2021). Long-term contextual effects in education: Schools and neighborhoods. *American Economic Journal: Economic Policy*, 13(2):336–77.

Li, D., Raymond, L. R., and Bergman, P. (2020). Hiring as exploration. Technical report, National Bureau of Economic Research.

Liang, A., Lu, J., and Mu, X. (2021). Algorithm design: A fairness-accuracy frontier. *arXiv preprint arXiv:2112.09975*.

Lundberg, S. J. (1991). The enforcement of equal opportunity laws under imperfect information: affirmative action and alternatives. *The Quarterly Journal of Economics*, 106(1):309–326.

Manski, C. F. (2004). Statistical treatment rules for heterogeneous populations. *Econometrica*, 72(4):1221–1246.

Manski, C. F. (2021). Econometrics for decision making: Building foundations sketched by haavelmo and wald. *Econometrica*, 89(6):2827–2853.

Massey, D. S. and Denton, N. A. (1993). *American Apartheid: Segregation and the Making of the Underclass*. Harvard University Press, Cambridge, MA.

Mazumder, B. (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.

Mogstad, M., Romano, J. P., Shaikh, A. M., and Wilhelm, D. (2023). Inference for ranks with applications to mobility across neighbourhoods and academic achievement across countries. *Review of Economic Studies*, page rdad006.

Mogstad, M. and Torsvik, G. (2021). Family background, neighborhoods and intergenerational mobility.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the us labor market. *The Quarterly Journal of Economics*, 118(2):549–599.

Munshi, K. (2020). Social networks and migration. *Annual Review of Economics*, 12:503–524.

Obermeyer, Z., Powers, B., Vogeli, C., and Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464):447–453.

Pope, D. G. and Sydnor, J. R. (2011). Implementing anti-discrimination policies in statistical profiling models. *American Economic Journal: Economic Policy*, 3(3):206–231.

Rambachan, A., Kleinberg, J., Mullainathan, S., and Ludwig, J. (2020). An economic approach to regulating algorithms. Technical report, National Bureau of Economic Research.

Rose, E. K., Schellenberg, J. T., and Shem-Tov, Y. (2022). The effects of teacher quality on adult criminal justice contact. Technical report, National Bureau of Economic Research.

Savage, L. J. (1954). *The Foundations of Statistics*. John Wiley & Sons.

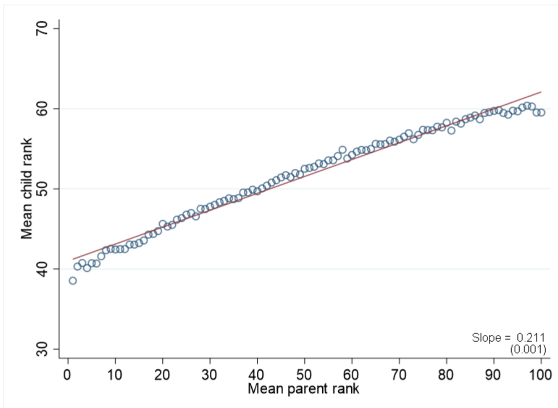
Theil, H. (1972). Statistical decomposition analysis with application in the social and administrative sciences.

Wald, A. (1950). Statistical decision functions.

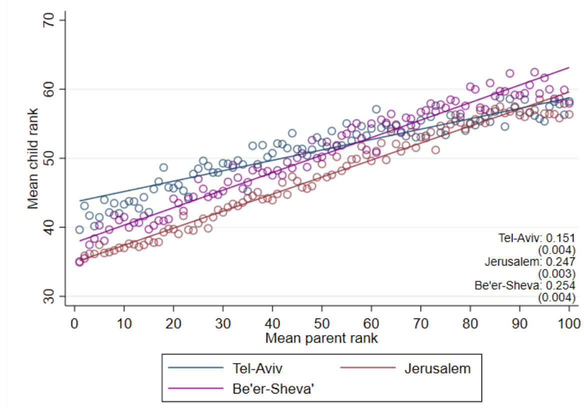
Wilson, W. J. (1987). *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. University of Chicago Press, Chicago.

A Additional Figures and Tables

Figure A.1: Relationship between parental income rank and child mean income rank at ages 28-30



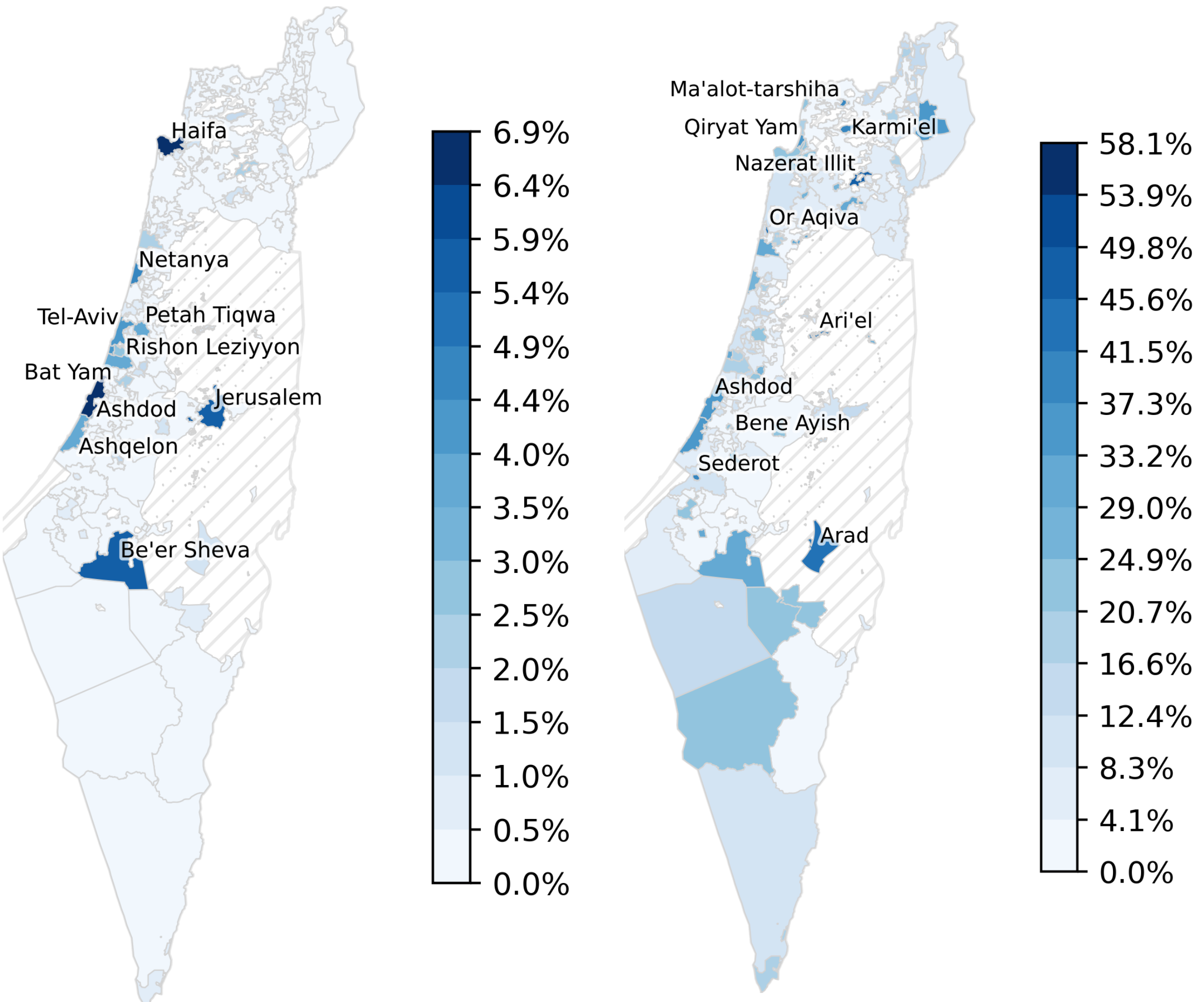
(a) Full population



(b) In Tel-Aviv, Be'er-Sheva, & Jerusalem

Note: This figure displays the relationship between parental income rank and mean child income rank at ages 28-30. Panel (a) displays the relationship in the full population, and panel (b) displays the relationship among children who lived in Tel-Aviv, Be'er-Sheva' and Jerusalem from birth to age 18.

Figure A.2: Immigrants' spatial distribution

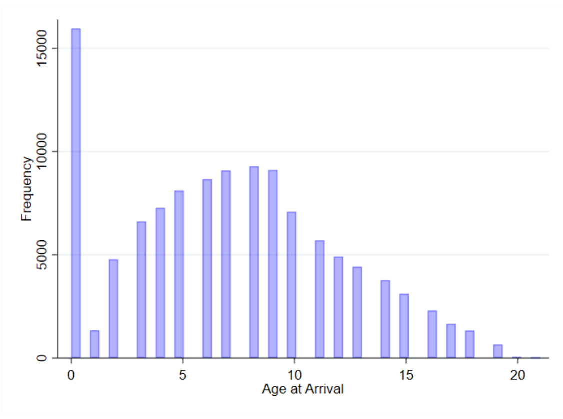


(a) Across city share

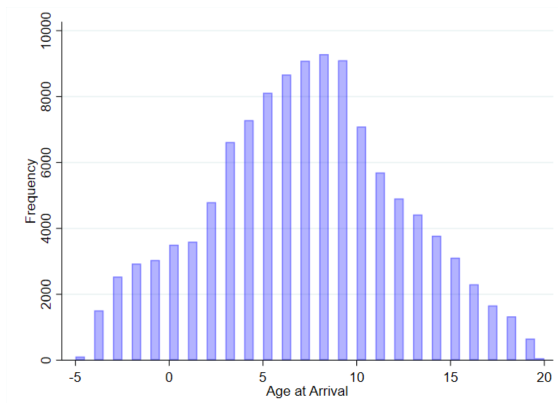
(b) Within city share

Note: This map presents the geographic distribution of immigrants across Israel. Panel (a) maps the share of immigrants in each location out of the immigrant population. Panel (b) maps the share of immigrants out of the whole population within each city. Location names are attached to the cities with the ten largest values. The values are grouped into 15 equally sized bins and collared accordingly. Source: The annual Local Authorities in Israel report of the Central Bureau of Statistics 2003.

Figure A.3: Distribution of children's age at arrival to Israel, Soviet immigrants



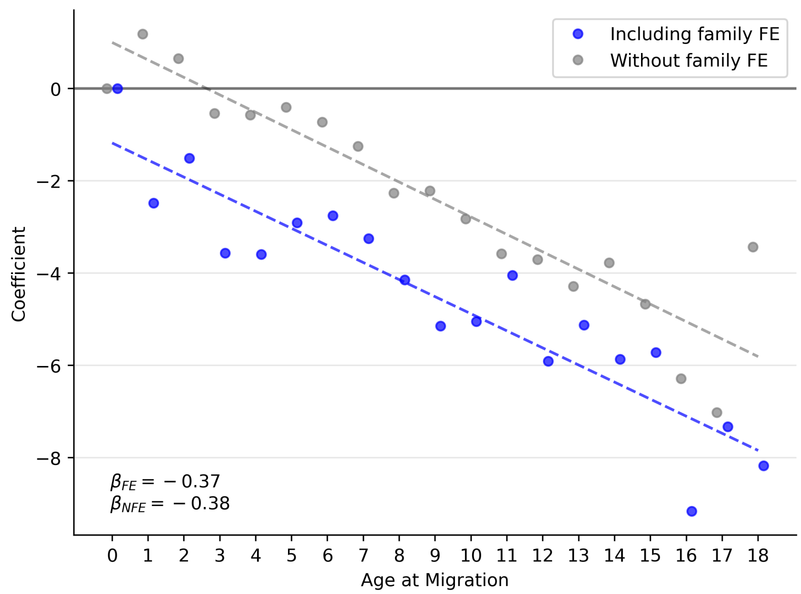
(a) Grouping the children who were born in Israel



(b) Child age at migration

Note: This figure presents the distribution of child's age at arrival in Israel among the Soviet immigrants entering Israel between the years 1989 to 2000, who were born between 1980-1995. Age zero refers to children whose parents immigrated to Israel between the year 1989 to 2000, and the child was born in Israel. In panel (a), we group all the children who were born in Israel in $age = 0$. In panel (b), negative age describes how many years after immigration the child was born.

Figure A.4: The effect of migrating to Israel by age

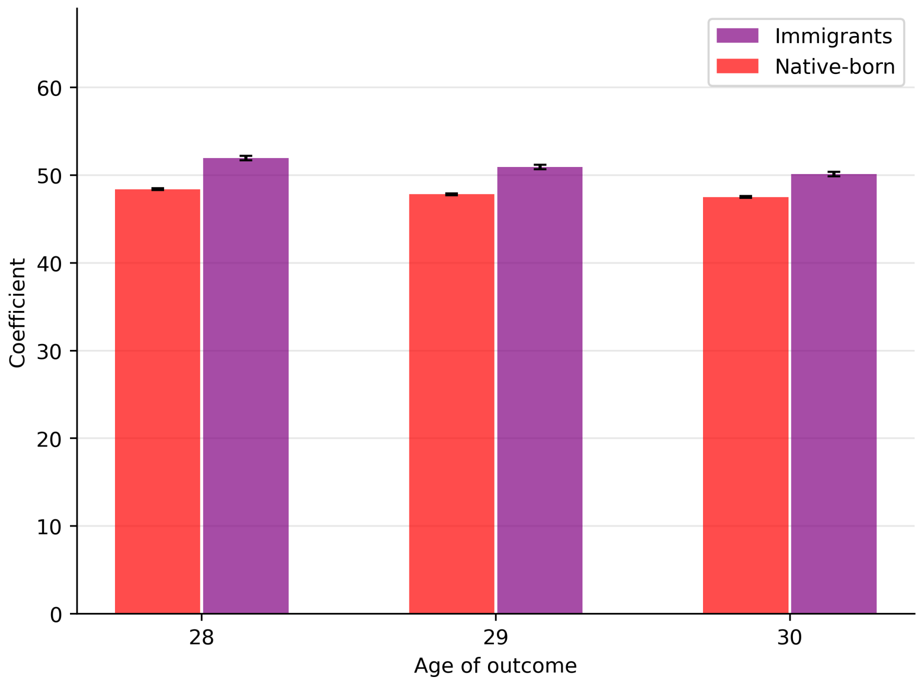


Note: This figure presents the relationship between the age at migration to Israel from the Former Soviet Union and children's earnings rank measured at age 28. Each dot represents the coefficient β_a, the effect of migrating to Israel at different ages, relative to age 0 (born in Israel)), estimated with the following regression:

$$y_i = \alpha + \sum_{a=1}^{18} \beta_a \mathbb{1}\{a(i) = a\} + \gamma_{f(i)} + \varepsilon_i,$$

where $a(i)$ is the migration age of child i, and $\gamma_{f(i)}$ is family fixed effect. In dashed lines are the linear fitted lines, with the corresponding slopes presented at the bottom. In blue, we present the results from the regression above with family fixed effects and in grey without.

Figure A.5: The expected rank of low-income children by immigration group and the age of outcome measurement

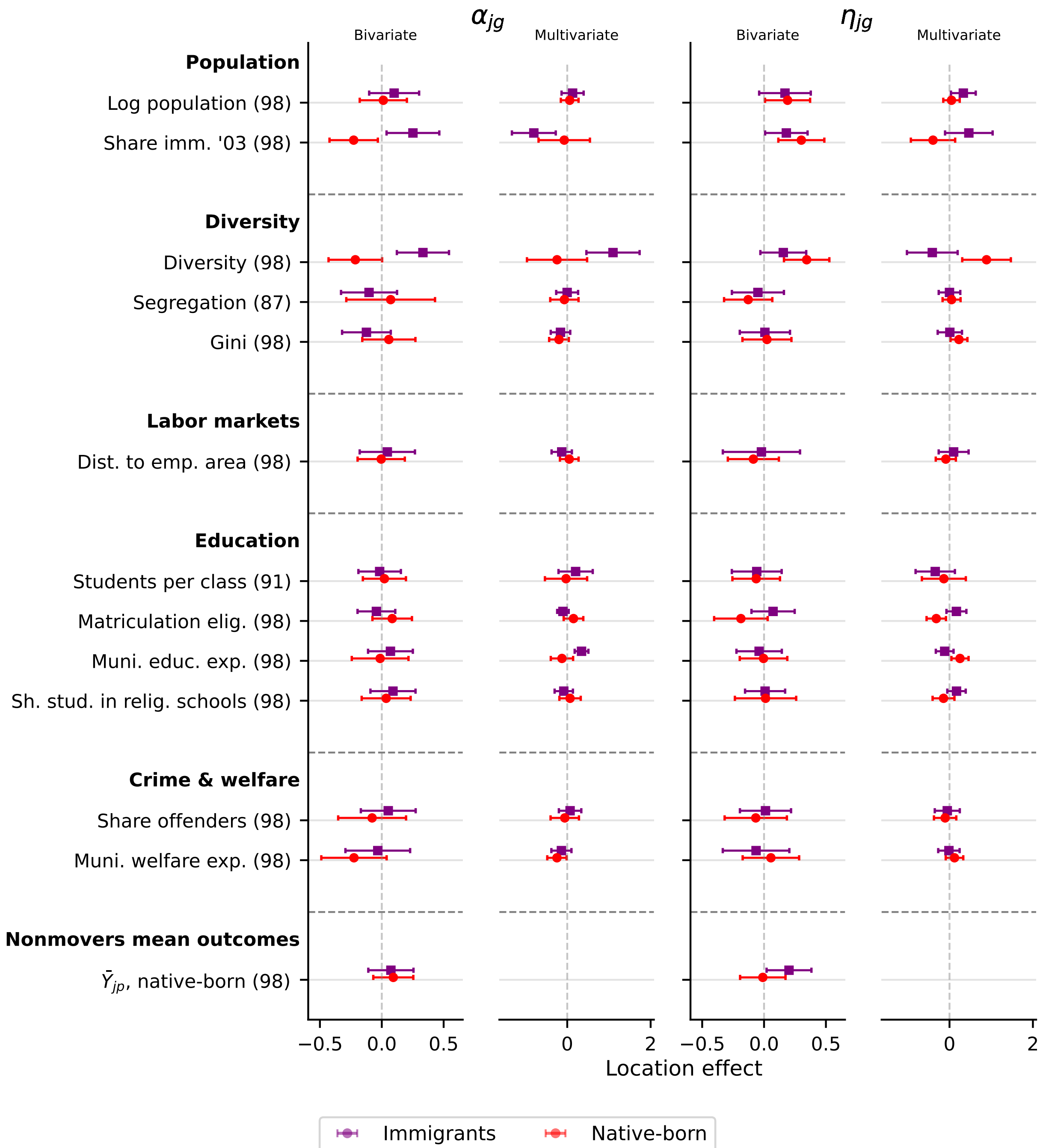


Note: This figure presents the expected rank of children of low-income parents, predicted using the intergenerational rank-rank regression estimates:

$$Y_{i,a} = \alpha + \beta p(i) + \varepsilon_i,$$

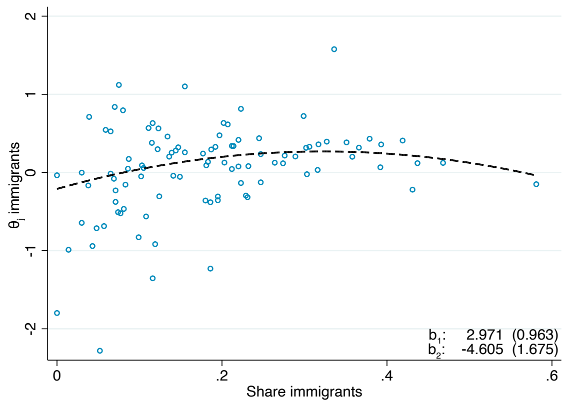
where $Y_{i,a}$ is the child i's earnings rank at age a, and $p(i)$ is child i's parents' rank. The bars display the predicted income percentile of a child with parents at the 25th percentile, $E[Y_{i,a}|p(i) = 25] = \hat{\alpha} + 25 \times \hat{\beta}$ by child's age. Confidence intervals are based on robust standard errors.

Figure A.6: Relationship between Location Effects and City Characteristics (α_{jg} and η_{jg})

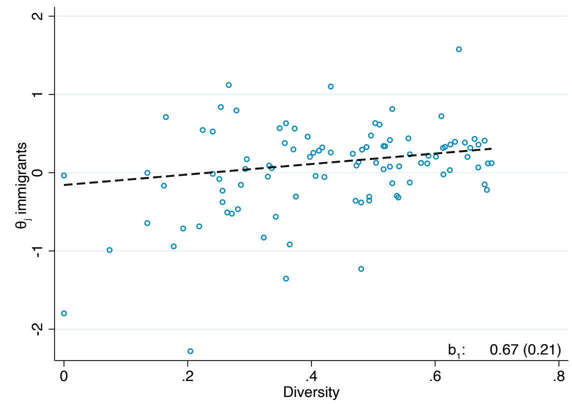


Note: This figure plots the relationship between city-level covariates and the parameters governing the location effects, the intercept α (left panel), and the slope η (right panel) for immigrants and natives. Each relationship is estimated with a feasible generalized least squares regression, reweighting the observations by the inverse of the Cholesky decomposition matrix of Σ, the variance of the estimated location effects, and with the location effects as the outcomes. Covariates are standardized to have a mean of zero and a standard deviation of one in the sample. In each panel, the first column plots the coefficients from regressions of effects on each covariate alone, and the second column plots the coefficients of a multivariate regression with all the characteristics simultaneously. Bars indicate 95% confidence intervals based on robust standard errors. Appendix Section B.3 provides a complete description of covariates definitions. The number of cities in each regression is in parentheses. Cases with fewer localities than the full sample (98) are due to missing values, or in the case of segregation, where values cannot be calculated for cities that do not have sub-areas (see Appendix Section H.1).

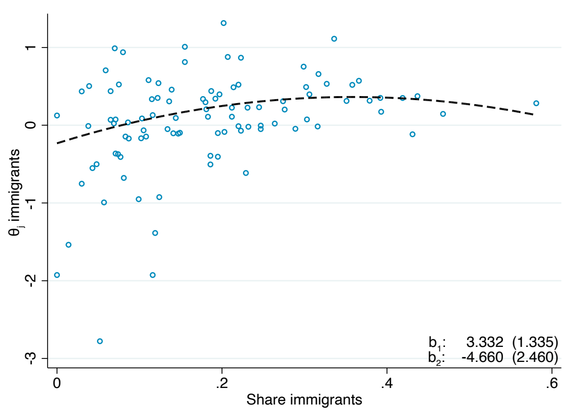
Figure A.7: Relationship between immigrant location effects and local demographics



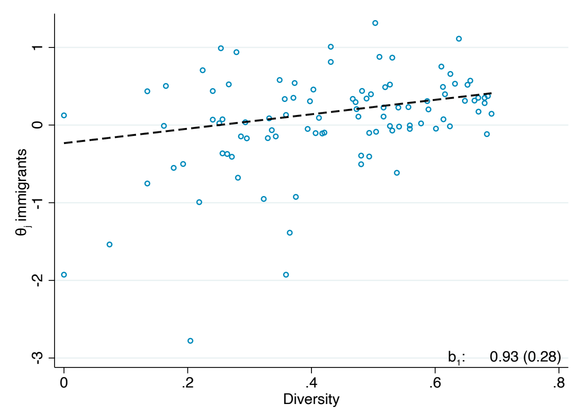
(a) θ_{25}, immigrant share



(b) θ_{25}, diversity



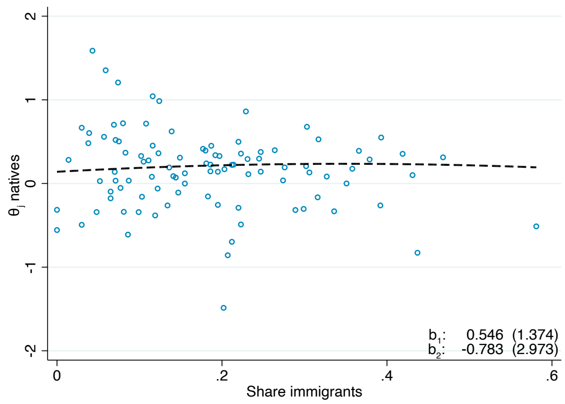
(c) θ_{75}, immigrant share



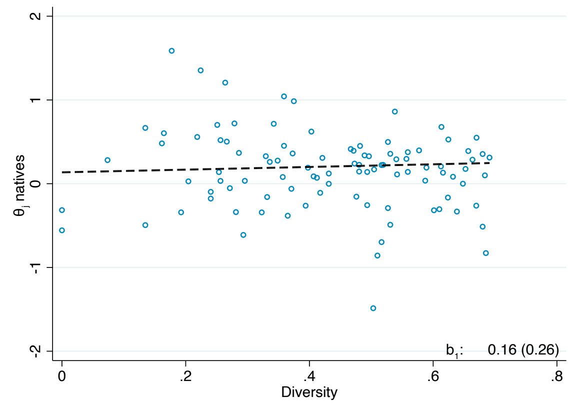
(d) θ_{75}, diversity

Note: This figure displays the relationship between immigrant location effects and the share of immigrants in the city, as measured in 2003. Panels (a) and (c) plot the location effects of families at the 25th and 75th percentile of the national income distribution against the share of immigrants in the city, where the dashed line is the quadratic function that best fits the data. Panels (b) and (d) plot the location effects of families at the 25th and 75th percentile of the national income distribution against the city diversity, where the dashed line is the linear regression line. Diversity is calculated as the log entropy index: $-(\pi_{\mathcal{I}} \ln(\pi_{\mathcal{I}}) + (1 - \pi_{\mathcal{I}}) \ln(1 - \pi_{\mathcal{I}}))$, where $\pi_{\mathcal{I}}$ is the share of immigrants in the city. All regressions are weighted by the number of observations.

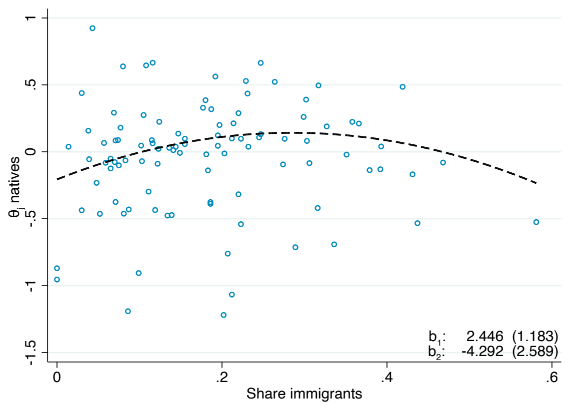
Figure A.8: Relationship between native-born location effects and local demographics



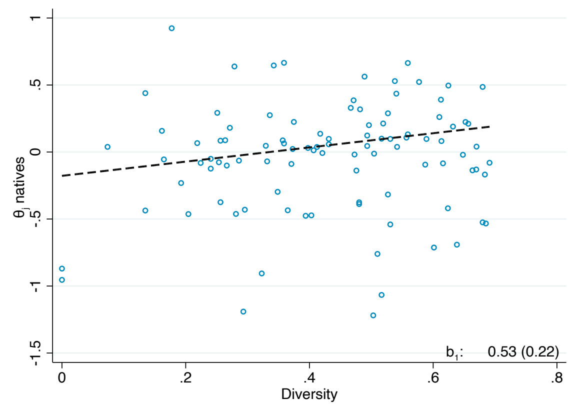
(a) θ_{25}, immigrant share



(b) θ_{25}, diversity



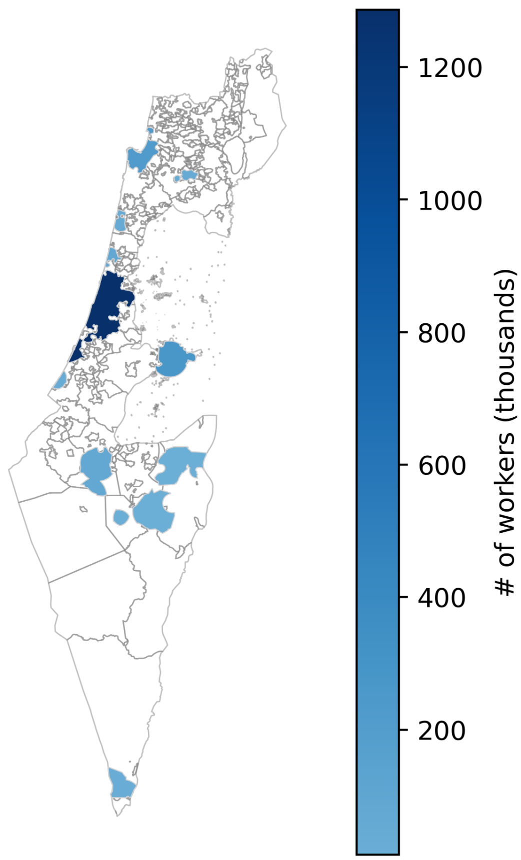
(c) θ_{75}, immigrant share



(d) θ_{75}, diversity

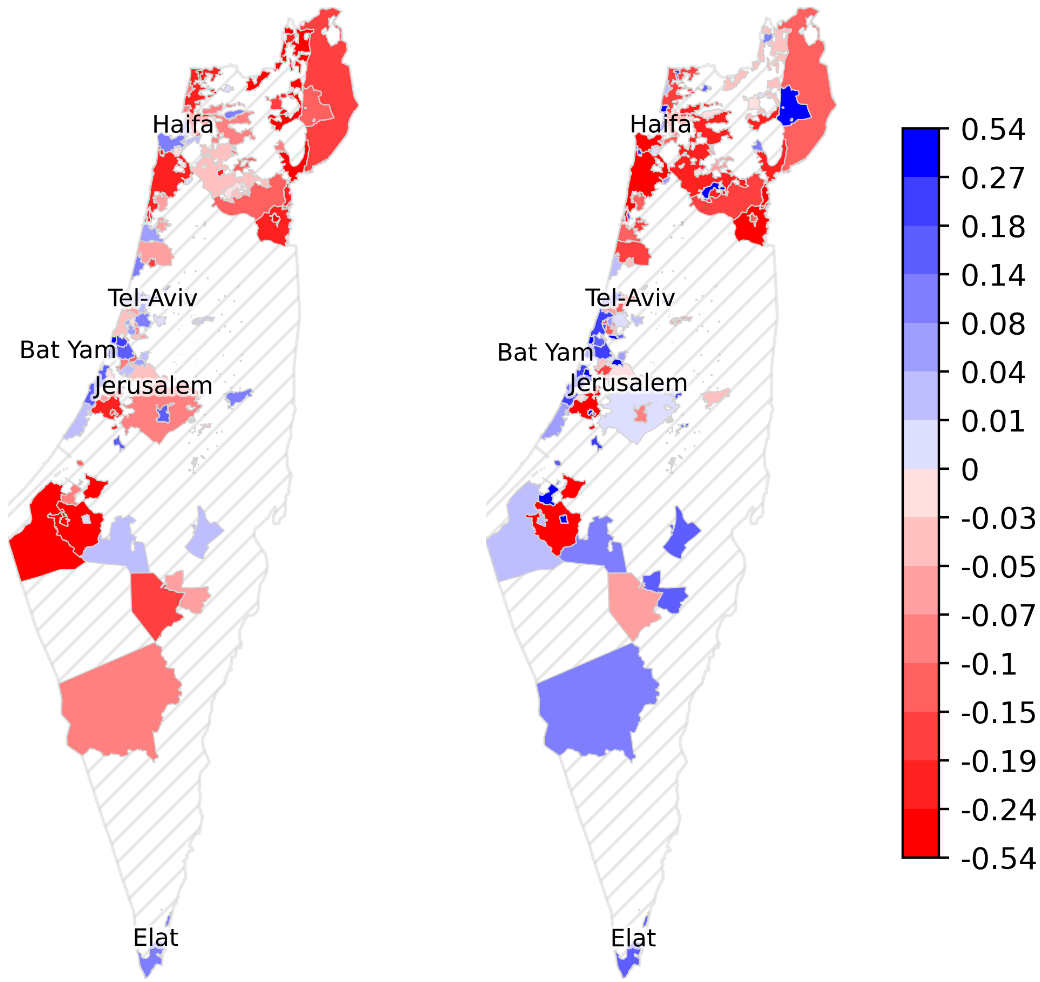
Note: This figure displays the relationship between native-born location effects and the share of immigrants in the city, as measured in 2003. Panels (a) and (c) plot the location effects of families at the 25th and 75th percentile of the national income distribution against the share of immigrants in the city, where the dashed line is the quadratic function that best fits the data. Panels (b) and (d) plot the location effects of families at the 25th and 75th percentile of the national income distribution against the city diversity, where the dashed line is the linear regression line. Diversity is calculated as the log entropy index: $-(\pi_{\mathcal{I}} \ln(\pi_{\mathcal{I}}) + (1-\pi_{\mathcal{I}})\ln(1-\pi_{\mathcal{I}}))$, where $\pi_{\mathcal{I}}$ is the share of immigrants in the city. All regressions are weighted by the number of observations.

Figure A.9: Employment center, census 2008



Note: This figure displays the number of workers (in thousands) in the major employment centers from the 2008 census.

Figure A.10: Posterior mean location effects, high-income families ($p = 75$)

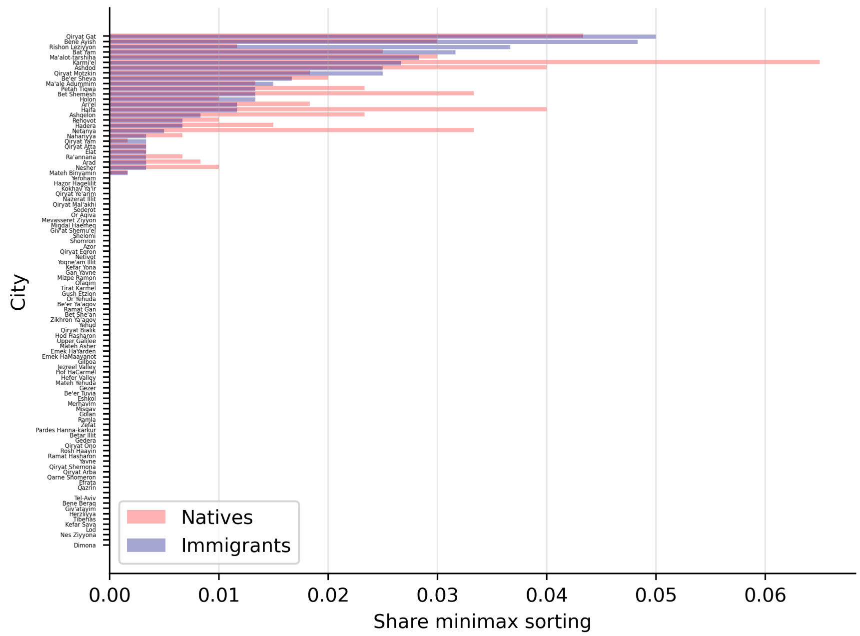


(a) Natives, θ_{75}

(b) Immigrants, θ_{75}

Note: These maps plot the children's posterior mean effect of year-long exposure to cities and regional councils in Israel on children's income rank at age 30 for immigrants and native-born children. Figures (a) and (b) display the location effects of low-income families whose parents are at the 25th percentile of the national income distribution, and figures (c) and (d) display the location effects of high-income families whose parents are at the 75th percentile of the income distribution. The maps are constructed by grouping cities into 15 equally sized groups where the more blue the area, the greater its effect from the mean, and the more red the area, the smaller the effect compared to the mean.

Figure A.11: The distribution of the least favorable sorting patterns



Note: This figure plots the probability that a housing voucher recipient family chooses to move to each Israeli city if families follow the minimax strategy and are facing the optimal policy of the minimax decisionmaker. Red bars display the sorting probabilities of native families, and blue bars display the sorting probabilities of immigrant families.

Table A.1: Standard deviation (diagonal) and correlation (off-diagonal) of the intercept and slope $(\theta_{jI}, \eta_{jI}, \theta_{jL}, \eta_{jL})$ location effects

	Natives		Immigrants	
	Cons. (1)	Rank-parents (2)	Cons. (3)	Rank-parents (4)
Natives				
Cons.	0.210			
	(0.053)			
Rank-parents	-0.696	0.002		
	(0.190)	(0.001)		
	[-0.884, -0.175]			
Immigrants				
Cons.	-0.198	0.113	0.228	
	(0.279)	(0.267)	(0.042)	
	[-0.710, 0.290]	[-0.362, 0.612]		
Rank-parents	-0.118	0.873	0.189	0.002
	(0.342)	(0.324)	(0.271)	(0.000)
	[-0.702, 0.435]	[0.448, 1.405]	[-0.183, 0.812]	

Note: This table reports the standard deviation in diagonal and correlation in off-diagonal of the vector of intercepts and slopes of the location effects of immigrants and native-born Israeli children. All variance components are weighted by the total number of city residents. Standard errors of the variance and covariances are based on the asymptotic variance, assuming location effects are drawn from a normal distribution. Standard errors of the correlations and standard deviations are calculated using the delta method. Square brackets display parametric bootstrapped equal-tailed confidence intervals.

Table A.2: Variance components and correlations, robustness to school fixed effects

	Baseline (1)	w/ school FE (2)	% explained variance (3)
(i) Low-income families (θ_{25})			
Std. Natives	0.173	0.128	0.453
	(0.050)	(0.059)	
Std. Immigrants	0.205	0.170	0.312
	(0.026)	(0.055)	
Std. Difference	0.319	0.545	
	(0.053)	(0.155)	
Immigrants-native corr	-0.081	-0.154	
	(0.306)	(0.673)	
	[-0.622, 0.457]	[-1.101, 0.711]	
(ii) High-income families (θ_{75})			
Std. Natives	0.198	0.157	0.371
	(0.041)	(0.045)	
Std. Immigrants	0.267	0.189	0.499
	(0.026)	(0.057)	
Std. Difference	0.307	0.474	
	(0.053)	(0.165)	
Immigrants-native corr	0.468	-0.076	
	(0.217)	(0.476)	
	[0.025, 0.879]	[-0.835, 0.673]	

Note: This table reports the bias-corrected variance components of year-long exposure location effects of native-born and immigrants from high- and low-income families. Column (1) reports our baseline estimates of location effects. Column (2) reports the variance components from a model with high school fixed effects. Column (3) reports the share of variance explained by high school fixed effects. Standard errors are calculated via the delta method.

Table A.3: Variance components and correlations, robustness to neighborhood reweighting

	P_{25}		P_{75}	
	Std.	χ^2 test $H_0 : \theta_j = \theta_1 \forall j$	Std.	χ^2 test $H_0 : \theta_{jI} - \theta_{jN} = c\ \forall j$
	(1)	(2)	(3)	(4)
Natives	0.9563	189.4	0.8924	202.1
	(0.2174)	[0.0000]	(0.1959)	[0.0000]
Immigrants	0.5040	464.0	0.5360	286.7
	(0.0749)	[0.0000]	(0.0559)	[0.0000]
Immigrants-native corr	0.0556		0.3098	
	(0.2185)		(0.1618)	

Note: This table reports the bias-corrected variance components of year-long exposure location effects of native-born and immigrants from high- and low-income families. Location effects were estimated by exploiting differences in the exposure time to cities in Israel using families who moved between cities and for immigrants exploiting the variation in the age of arrival to the country. This table reports the robustness exercises reweighting the regression by the origin-destination number of observations. Columns (1) and (2) present the results for families from the 25th percentile of the income distribution, and columns (3) and (4) present the results for families from the 75th percentile of the income distribution. The first two rows in columns (1) and (3) display the standard deviation of location effects, while the third row presents the correlation between natives' and immigrants' location effects estimate as the covariance divided by the standard deviations. Standard errors are calculated via delta method.

Table A.4: Location effect hyperparameters

	Low-income families (θ_{25}) (1)	High-income families (θ_{25}) (2)
(i) Intercept		
$\beta_{\mathcal{N}}$	0.423	-0.239
	(0.184)	(0.185)
$\beta_{\mathcal{I}}$	-0.257	-0.305
	(0.149)	(0.175)
(ii) Diversity		
$\beta_{\mathcal{N}}$	0.098	0.442
	(0.250)	(0.196)
$\beta_{\mathcal{I}}$	0.477	0.726
	(0.212)	(0.268)
(iii) Welf. expnd.		
$\beta_{\mathcal{N}}$	-0.063	-0.041
	(0.013)	(0.012)
$\beta_{\mathcal{I}}$	-0.003	-0.009
	(0.015)	(0.020)
(iv) Log SE		
$\beta_{\mathcal{N}}$	-0.031	-0.198
	(0.081)	(0.075)
$\beta_{\mathcal{I}}$	-0.106	-0.115
	(0.059)	(0.086)
(v) Std. devs. and correlation		
$\sigma_{\mathcal{N}}$	0.125	0.089
	(0.048)	(0.044)
$\sigma_{\mathcal{I}}$	0.201	0.256
	(0.042)	(0.045)
ρ	-0.381	0.123
	(0.935)	(0.987)
(vi) Implied total std. and correlation		
Natives	0.189	0.178
	(0.041)	(0.035)
Immigrants	0.243	0.314
	(0.034)	(0.036)
Correlation	-0.123	0.375
	(0.249)	(0.224)
	[-0.581, 0.363]	[0.024, 0.870]
R^2 Natives	0.563	0.750
R^2 Immigrants	0.316	0.335
# of cities	98	98

Note: This table reports the estimated parameters and standard errors of the joint distribution of native-born and immigrant location effects, assuming they are drawn from a normal distribution. Column (1) presents the estimates for low-income families, θ_{25}, and column (2) for high-income families θ_{75}. Panels (i)-(iv) report the mean native (subscript N) and immigrant (subscript I) location effects, which is a linear function of city-level diversity, municipality welfare expenditure per capita, and log standard error. Panel (v) reports the standard deviation (σ) of natives and immigrants and the correlation (ρ) of the random effect, and panel (vi) reports the implied total standard deviation of the location effects reported, which is defined as $\sqrt{\beta_g^2 \mathbb{V}(z) + \sigma_g^2}$ for every group $g \in \{\mathcal{I}, \mathcal{N}\}$, and the implied correlation, which is the ratio between $\mathbb{C}ov(z_j'\beta_{\mathcal{I}}, z_j'\beta_{\mathcal{I}}) + \rho\sigma_I\sigma_N$ and the product of the implied standard deviation of immigrants and natives. R^2 is defined as the group specific ratio between variance share $\frac{\beta_g^2 \mathbb{V}(z)}{\beta_g^2 \mathbb{V}(z) + \sigma_g^2}$. Panels (i)-(iv) were estimated by a city-size weighted least regression, and panel (v) was estimated via the city-size weighted method of moments. In panels (i)-(iv), robust standard errors are reported in parenthesis, and in panels (v)-(vi), parentheses report the parametric bootstrapped standard errors, and square brackets report parametric bootstrapped equal-tailed confidence intervals.

Table A.5: Top and bottom cities based on the long-run location effects, high-income families

(A) Sorted by natives' loc. effects				(B) Sorted by immigrants' loc. effects			
Loc. name	Posterior mean imm. (1)	Posterior mean natives (2)	Share immigrants (3)	Loc. name	Posterior mean imm. (4)	Posterior mean native (5)	Share immigrants (6)
(i) Top 10							
Bat Yam	0.338	0.267	0.317	Akko	0.542	-0.076	0.277
Holon	0.145	0.179	0.177	Or Yehuda	0.466	-0.031	0.155
Qiryat Gat	0.267	0.168	0.327	Ofaqim	0.413	0.001	0.299
Bet Shemesh	-0.098	0.166	0.247	Netivot	0.374	-0.084	0.223
Rishon Leziyyon	0.195	0.150	0.187	Shomron	0.362	0.007	0.080
Ashdod	0.207	0.136	0.366	Qazrin	0.342	-0.148	0.336
Elat	0.173	0.129	0.180	Bat Yam	0.338	0.267	0.317
Qiryat Motzkin	-0.019	0.128	0.220	Qarne Shomeron	0.316	-0.045	0.155
Kokhav Ya'ir	0.000	0.089	0.186	Qiryat Mal'akhi	0.287	-0.155	0.220
Ma'ale Adummim	-0.035	0.088	0.149	Ramla	0.286	0.050	0.269
Petah Tiqwa	0.006	0.087	0.212	Yavne	0.281	-0.015	0.123
	⋮	⋮	⋮		⋮	⋮	⋮
(ii) Bottom 10							
Merhavim	-0.479	-0.542	0.000	Merhavim	-0.479	-0.542	0.000
Shelomi	0.226	-0.509	0.202	Gedera	-0.410	-0.161	0.116
Qiryat Arba	-0.064	-0.382	0.186	Emek HaMaayanot	-0.365	-0.224	0.052
Bet She'an	-0.139	-0.355	0.074	Be'er Tuvia	-0.351	-0.196	0.014
Upper Galilee	-0.030	-0.341	0.134	Hof HaCarmel	-0.258	-0.215	0.119
Emek HaYarden	-0.220	-0.333	0.099	Haifa	-0.253	0.086	0.272
Eshkol	0.021	-0.270	0.075	Misgav	-0.225	-0.096	0.044
Hazor Hagelilit	0.007	-0.237	0.116	Emek HaYarden	-0.220	-0.333	0.099
Mateh Asher	-0.164	-0.224	0.000	Qiryat Eqron	-0.200	-0.081	0.124
Emek HaMaayanot	-0.365	-0.224	0.052	Jezreel Valley	-0.196	-0.034	0.062

Note: This table presents the top 10 and bottom 10 locations in Israel with respect to the long-run effects on children's income rank at age 28 whose parents are from the 75th percentile of the income distribution. Panel (A) is sorted based on the posterior mean location effects of native-born children, and panel (B) is sorted based on the posterior mean location effects of immigrants. Columns 1-2 and 4-5 present the posterior mean location effect, and columns 3 and 6 present the share of former Soviet Union immigrants.

Table A.6: Top selected Israeli cities, $K = 10$, status-quo sorting normalization

	Psterior mean					
	Native-born	Immigrants	Average	$\mathbb{E}[\min\{\vartheta_{\mathcal{N}}, \vartheta_{\mathcal{I}}\}\|\mathcal{Y}]$	Worse than status-quo	Selected by maximin ($\mathcal{I}/\mathcal{N}$/city)
Loc. name	(1)	(2)	(3)	(4)	(5)	(6)
Qiryat Gat	356.6	526.1	470.7	312.5	0.023	Yes (93.4)
Karmi'el	270.5	475.7	395.1	233.1	0.072	Yes (93.4)
Rishon Leziyyon	236.0	479.6	434.0	226.4	0.033	Yes (93.4)
Ma'alot-tarshiha	348.7	348.7	348.7	220.6	0.083	Yes (93.4)
Bat Yam	211.9	579.4	462.9	209.9	0.046	Yes (93.4)
Yavne	394.2	253.2	270.5	183.2	0.111	Yes (93.4)
Mateh Binyamin	348.3	271.5	280.4	181.9	0.109	Yes (93.4)
Ashqelon	265.3	261.7	263.0	173.9	0.075	Yes (93.4)
Ramla	296.9	248.3	261.4	171.6	0.093	Yes (93.4)
Arad	347.1	258.2	295.4	170.9	0.129	Yes (93.4)
Ra'annana	211.2	367.1	345.9	157.0	0.162	
Holon	160.4	471.1	416.1	155.3	0.113	
Be'er Sheva	173.4	281.1	248.2	136.7	0.091	
Qarne Shomeron	270.2	316.3	309.1	136.4	0.214	
Mevasseret Ziyyon	243.0	288.2	283.2	119.3	0.230	

Note: This table reports the list of the top 15 Israeli cities sorted by the within-city posterior immigrant-native minimum location effect, where the location effects of both groups are normalized in comparison to the expected effect under the status-quo sorting patterns of each group. Location effects are in Shekel value (1 US \$ $\approx$ 3.4 ILS) and represent earnings returns at age 28 from spending one year in city j, compared to the average returns under the status-quo sorting patterns. Columns 1-3 report the posterior mean of native-born children, immigrants, and the average. Column 4 reports the posterior minimum location effect across immigrants and natives. Column 5 reports the posterior probability that the location effects of immigrants or the location effects of native-born are lower than the average effect under the status quo sorting patterns. Column 6 reports which cities are selected as the top 10 cities based on the minimum (N/I/city) policy that ranks lists of 10 cities based on their posterior minimum location effect across all cities and groups, where the posterior minimum of the selected list is presented in parentheses.

B Data and Definitions

B.1 Data construction

This appendix provides a general overview of the data construction and restrictions. Our data construction starts with base demographics data, which contains the entire population of Israel born between 1950 and 1995, their years of birth and death, fake identifiers of individuals and their parents, and the country of birth of both parents and the child. To this file, we merge the annual population registry files which contain the locality of residence at the city and statistical area level of children and both parents separately, the immigration year and country of origin of the child and each parent, and tax records files for the years 1995, 2001, and 2002 to 2019. In addition, we supplement the data with parental incomes available starting from 1983. We correct all income values for inflation to prices of 2016, sum the total earnings from all sources (employed or self-employed earnings), and define the ranks of children and parents within the income year relative to the entire sample population. To define the immigrants for our analysis, we follow the rule that if at least one parent was born in the USSR and immigrated starting from 1989, we define the child as an immigrant. In this process, we drop anomaly observations for what we believe are data input mistakes in the administrative records, which include individuals with birth dates after their death dates ($\approx$ 450 observations), very early birth years of parents, below 1950 (10 observations), non-matching single parent identifiers to tax records (113 observations), and negative parental earnings found only in the year 1983 ($\approx$ 950 observations).

To this file, we merge school identifiers and localities at the city level annually for the period 1995-2016. Next, we construct the geographic mobility variables. We use the registry data as the main source for locality information at the annual level. Since we observe both the mother's, father's, and child's locality each year, we define the single series of locations as the parent's location for the parent who shares the most years at the same location with the child. We define a move in every year in which the locality in the series is different from the locality in the previous year. Since the registry data is missing the years 1996-1999 and 2001, we use the school locality information to accurately pinpoint any move that happens in the missing years' span. That is, if we identify a change in locations during missing locality information, we use the school locality data, in cases that it matched either the origin or the destination, to establish the correct move age.

After defining for every child the origins, destinations, and move ages for all moves, we count the number of children that are associated with origin and destination before the age of 18 and restrict attention to localities that have at least 5 children in a location as an origin, and at least 5 children in a location as a destination, with at least 45 children in the location at the total, regardless whether the location is an origin or destination. This results in 98 localities out of the total 152 localities in Israel that are in our sample.

B.2 Data comparison

In Table B.1, we compare our sample to those used in the analyses conducted in Chetty and Hendren (2018a) and Deutscher (2020). Panel (A) presents the sample sizes and cohort spans, while Panel (B) presents the period of variable availability of location and income information. We present both the sample sizes of non-movers and 1-time movers, as in the papers above, but also migrants who were born in Israel. It is important to note that our sample definitions are different for immigrants compared to locals. In the case of locals, as is the case in the literature in general, the sample of analysis is the population of children who move during their childhood. In the case of immigrants, we exploit the fact that they have arrived in Israel and settled in their residence location at different ages and measure their exposure time to a location starting from their arrival age. Therefore, since we define permanent residents as children who do not move *within* Israel, the immigrant children who are non-movers are those who are included in our analyses. In addition, children of immigrant families from the immigration waves of 1989-2000 who were born in Israel are also included in our sample with an age of arrival of 0. Therefore, the sample of analysis combining both permanent resident immigrants and local 1-time movers, sums up to 311 thousand observations, very similar to the sample size in Deutscher (2020).

Our sample cohorts are similar to those used in Deutscher (2020), with a span of 14 cohorts, compared to 9 in Chetty and Hendren (2018a). The period of years in which we observe outcomes for children is also comparable in size to Deutscher (2020), with 25 years, compared to 17 years in Chetty and Hendren (2018a). This allows us to observe outcomes starting from age 5 for natives, and 0 for immigrants. Since Israel's population is significantly smaller than that of the US and Australia, our sample is considerably smaller. Also note that our cohort restriction together with the fact that the immigration wave was during the years

1989-2000, implies that we have only a few observations with immigration age above 17.

Table B.1: Comparison of data with Deutscher (2020) and Chetty and Hendren (2018a)

	Deutscher (2020)	Chetty and Hendren (2018a)	This paper	
			Migrants	Native
	(1)	(2)	(3)	(4)
(A) Sample size and time span				
Birth cohorts	1978-1991 (14 years)	1980-1988 (9 years)	1980-1991 (12 years)	
Permanent residents	1,683,800	19,499,662	130,727	883,144
Movers	313,900	1,553,021	25,542	116,572
(B) Location and income information				
Data range	1991-2015 (25 years)	1996-2012 (17 years)	1990-2019 (30)	1990-2019 (30)
Potential range of age at move	1-39 years	9-32 years	0-29 years	5-29 years
Analysis range of age at move	2-34 years	9-30 years	0-17 years	5-29 years

Note: The first and second columns are based on Table 2 in Deutscher (2020). Columns (3) and (4) present the samples used in this paper for immigrants and natives, respectively.

B.3 City Level Data

In this Appendix, we provide detail on the locality-level variables used in Section **??**.

- **Gini index for inequality** calculated at the locality level, taken from the annual Local Authorities in Israel report of the Central Bureau of Statistics (1998), is measured based on the gross earnings of employees in 1998, using administrative records from The National Insurance Institute records.

- **The Theil (1972) index for segregation** constructed using the 2000 tax records data for earnings, combined with the population registry for the city of residence. The segregation index is calculated in the following way. For every group g[35] we denote π_g the share of individuals in a given city. Let $s = 1, .., N$ index the statistical areas in each city. Analogously, we calculate in every statistical area π_{gs}, the group share in statistical area s. For every city and every statistical area, we measure the entropy index $E = \sum_g \pi_g log\frac{1}{\pi_g}$, and $E_s = \sum_g \pi_{gs} log\frac{1}{\pi_{gs}}$. The degree of segregation in every city is defined as:

$$H = \sum_s (\frac{pop_s}{pop}\frac{E - E_s}{E})$$

[35] We calculate this variable for two sets of population groups: immigrants and non-immigrants, and the different ethnic groups in Israel.

where pop_s is the total population of statistical area s and pop denotes the total population of that city. The segregation index H ranges between 0 and 1. When it equals 1 there is no heterogeneity within statistical areas indicating a high level of segregation. When $H = 0$ then all the statistical areas in the city have the same group shares as the city as a whole. Note that we cannot calculate segregation values for localities for which we do not observe sub-cities or sub-localities.

- **Diversity**: similarly to the segregation index, we construct a diversity index based on the entropy index. However, instead of focusing on the discrepancy between the entropy of the city, denoted earlier as E, and that of its individual statistical areas (which results in a measure of segregation), we calculate the entropy directly at the city level. Specifically, for each city, we calculate its diversity as:

$$E_d = \sum_g \pi_g \log \frac{1}{\pi_g}$$

  In this equation, E_d represents the diversity index for each city, π_g is the proportion of group g in the city, and the summation is performed over all groups g in the city.

  The diversity index E_d ranges between 0 and $\log G$, where G is the number of groups. When $E_d = 0$, it means there is no diversity, indicating that the city is entirely composed of a single group. Conversely, when $E_d = \log G$, it means there is maximum diversity, indicating that each group is equally represented in the city.

  By constructing the diversity variable in this way, we are able to capture the variety and evenness of group representation in each city, which is the essence of diversity.

- **Share criminal offenders**: the proportion of individuals within the locality who have been charged with a serious crime for which the potential punishment is imprisonment, taken from Fogel (2006), based on police records of 2002.

- **Municipality welfare expenditure**: the total government expenditure on welfare payments per capita, taken from the annual Local Authorities in Israel report of the Central Bureau of Statistics (1998), based on administrative records from The National Insurance Institute records.

- **High-school Bagrut eligibility**: the proportion of 12th-grade students of the years 1999 and 2000 who were eligible for a Bagrut certificate, taken from the annual Local Authorities in Israel report of the Central Bureau of Statistics (2000), based on administrative records from the Ministry of Education.

- **Distance to an employment center**: we use the Central Bureau of Statistics definition of central employment hubs, identified using a spatial interpolation model (Inverse Distance Weighted) to locate the largest concentrations of employees. The processing was based on workplace data from the 2008 population census, and it results in about 11 main employment hubs in Israel: Tel-Aviv, Haifa, Jerusalem, Hadera, Netanya, Be'er Sheva, Ashkelon, Eilat, Nazareth, Dimona, and Arad. We calculate the distance of each locality center to the nearest employment hub border.

- **Peripherality index**: we use the peripherality score of Israeli localities based on their geographic proximity to major population centers. Developed by the Central Bureau of Statistics, it is based on factors like distance from markets, employment hubs, and the Tel Aviv district. The index averages "potential accessibility" which measures the ease of accessing opportunities from a location with a weight of $\frac{1}{3}$, and proximity to the border of Tel Aviv with a weight of $\frac{1}{3}$. Large localities' scores are averages of their sub-localities, adjusted for population size.

- Based on the 1995 census, we also use the following shares:
  - **Within-city immigrants shares**: the proportion of immigrants out of the city population, capturing the density of immigrants within a city.
  - **City share of immigrants**: the proportion of immigrants out of the immigrant population in Israel who live in the city, as a measure for the dispersion of immigrants across Israel.

C One Move vs. Multiple Moves

Our approach diverges from the traditional literature by including not just one-time movers but also those who move twice (Chetty and Hendren, 2018a,b). In this appendix section, we describe the migration patterns of immigrnats and natives between districts in Israel. We begin by examining the migration patterns visualized in our data, followed by a comparison of the origins and destinations between the two groups and between the first and the second moves. We then recreate the main results in the paper solely on the single move population as a robustness test.

Panel (a) of Figure C.1 presents the transitions between districts for native-born families in Israel. Panel (b) presents the same figure for immigrants. These plots provide a visual illustration of the moves that we exploit in our data. For natives, we use both the first and the second move within Israel, and for immigrants, we use the migration from the Former Soviet Union to localities in Israel as their first "move", and their moves within Israel after arrival as the second move. We note the large proportion of one-time movers compared to two-time movers, at roughly seven times larger than the sample of two-time movers. Second, we learn about the relative size of the transitions between districts. We note that the biggest origins and destinations for natives are the central ones: Tel-Aviv and the Center district, which includes the close periphery of Tel-Aviv. For migrants, the South district takes in a larger share of families in the first move, while in the second, the Center district is the biggest destination. Third, we learn that a large proportion of moves are of relatively short distances between locations within districts.

Figure C.2 presents the transition matrices between the districts of Israel, where each cell represents the proportion of children who make the move from the row location to the column location, out of all the movers in the matrix. We compare the first move of one-time movers in Panel (a) to the first move of two-time movers in Panel (b). We find that these populations exhibit very similar patterns, with close proportions throughout the tables. The main differences are the moves between Tel Aviv and the Center district. One-time movers seem to move more from Tel Aviv to the center, with 19% of all one-time movers making that particular move, which is the largest among this group, while only 12.4% of two-time movers make that move. The largest proportion move among the two time movers is within the Tel Aviv district.

Figure C.3 describes the second move. Comparing this table to the tables in Figure C.2, we find again a large similarity in the distribution of moves. The main difference is in the small proportion of moves from Tel Aviv to the Center district, at 8.2%, while the reverse move is somewhat higher than in previous tables, with 7.9% of movers from the Center district to

Tel Aviv district.

In Table C.3, we recreate the results from Table 2 of the across-city heterogeneity estimates, restricting the sample to the single move population. First, note that removing the two-time movers decreases the number of localities that are included in the analysis to 92 (from 98). Second, we also find that the standard deviations using the overlap cities are all qualitatively similar to, and have overlapping confidence intervals with those found in our main analysis in Table 2. Lastly, the standard errors of most estimates in the overlap cities' sample (right panel), are smaller in our main analysis using both types of movers.

Lastly, we estimate the within-city heterogeneity, as done in Table 3, on the one-time movers' sample. The results are presented in Table C.4. We find that the covariance structure of the location effects in the one-time movers sample and twice movers sample is qualitatively the same. We perform chi-squared tests, assessing the null hypothesis of equal distributions between groups across the combinations of the matrices above, and fail to reject the null that they are similar.

Tables C.1 and C.2 provide a description of the sample of immigrants and natives, accordingly. First, we note that values are similar when comparing the full sample of cities (left panels) and the analysis sample of 98 cities (right panels) for both groups. Among the immigrants, those who moved after arrival to Israel had higher parental earnings. This is the case for native parents as well, where one-time movers have higher incomes than both stayers and two-time movers. Comparing the two groups, we note that native children generally outearn their immigrant counterparts, although these differences are quite small. These small differences at the children's generation level are notable, especially given the gap in parental earnings, as native children's parents boast significantly higher incomes compared to the parents of immigrant children. When it comes to educational outcomes, native children are slightly more likely to have attained a Bachelor's degree by age 27 than immigrant children. Lastly, immigrant children tend to have much higher intermarriage rates, which primarily reflects their smaller proportion in the population.

Table C.1: Sample descriptive statistics, immigrants

	All cities			98 sample cities		
	Stayers	Movers	All	Stayers	Movers	All
	(1)	(2)	(3)	(4)	(5)	(6)
(A): Children						
Income at 28	66,926	67,847	67,108	68,111	68,536	68,191
Rank at 28	52.68	51.97	52.54	53.43	52.45	53.24
BA at 27	0.156	0.164	0.157	0.157	0.161	0.158
Inter-marriage rates	0.25	0.25	0.25	0.28	0.24	0.23
(B): Parents						
Parents income	125,859	152,698	131,670	124,521	150,317	129,997
Parents rank	45	48.11	45.7	45.2	47.9	44.8
Num. of children	125,959	30,310	156,269	112,472	26,192	138,664

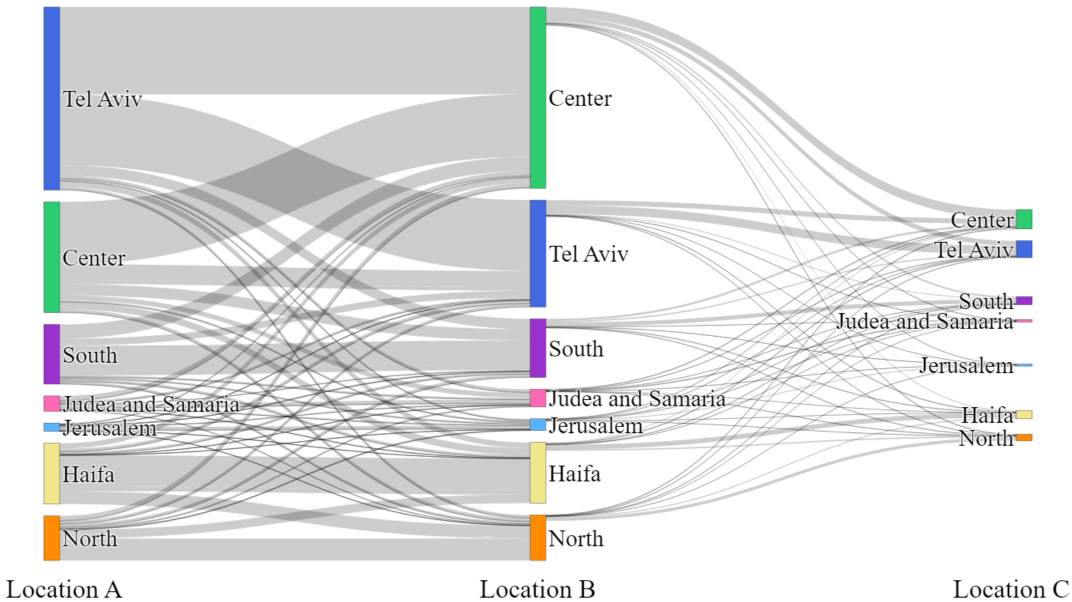
Note: This table presents the means of immigrant children (Panel A) and their parents (Panel B). We present the statistics for stayers, migrants who were either born in Israel or immigrated and did not move within Israel, and movers, who moved within Israel up to two times. In the left panel, we include all cities in our sample, and in the right, we include only the 98 cities for which we can estimate effects for both immigration groups and hence are included in our main analysis. Income variables are measured in Israeli Shekels (≈ 3.4$). Inter-marriage rates refer to the portion of immigrants who marry a native-born.

Table C.2: Sample descriptive statistics, natives

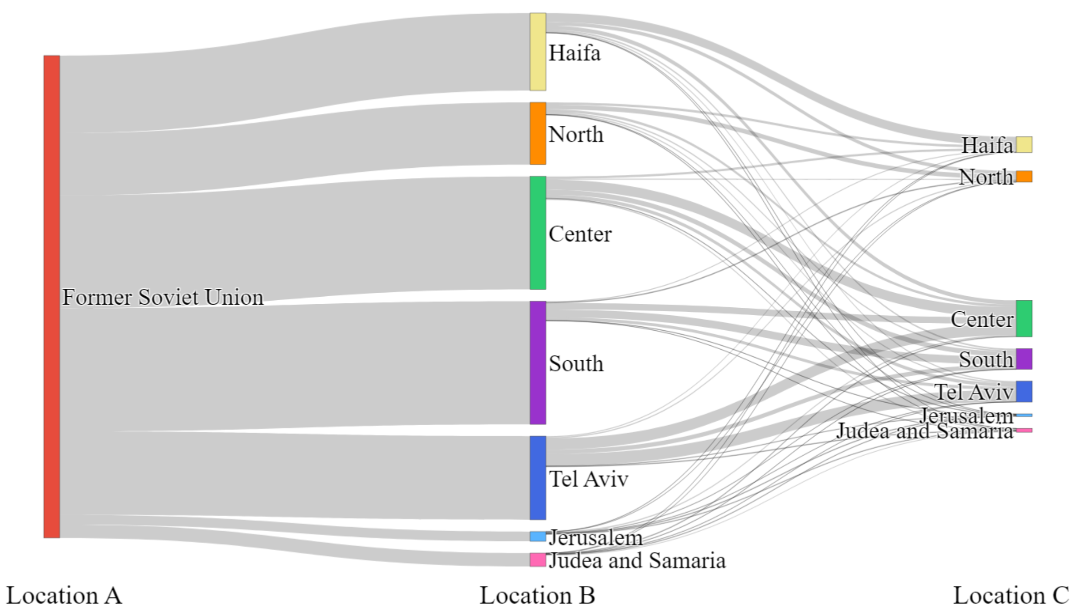
	All cities			98 sample cities		
	1 move	2 moves	Stayers	1 move	2 moves	Stayers
	(1)	(2)	(3)	(4)	(5)	(6)
(A): Children						
Income at 28	70,951	69,196	66,754	71,964	69,757	68,266
Rank at 28	53.38	51.06	53.09	53.86	51.36	53.86
BA at 27	0.21	0.17	0.18	0.21	0.17	0.19
Inter-marriage	0.049	0.053	0.043	0.049	0.051	0.042
(B): Parents						
Parents income	242,214	198,267	199,978	239,465	194,519	201,033
Parents rank	64.1	58.4	58.9	64.0	57.9	59.3
Num. obs	101,562	13,758	610,945	83,346	11,260	492,104

Note: This table presents the means of native-born children (Panel A) and their parents (Panel B). We present the statistics for one and two-time movers up to age 18, as well as stayers, natives who did not move within Israel since birth and up to age 18. In the left panel, we include all cities in our sample, and in the right, we include only the 98 cities for which we can estimate the effects for both immigration groups and hence are included in our main analysis. Income variables are measured in Israeli Shekels (≈ 3.4$). Inter-marriage rates refer to the portion of natives who marry an immigrant.

Figure C.1: Families flows in first and second move



(a) Native-born

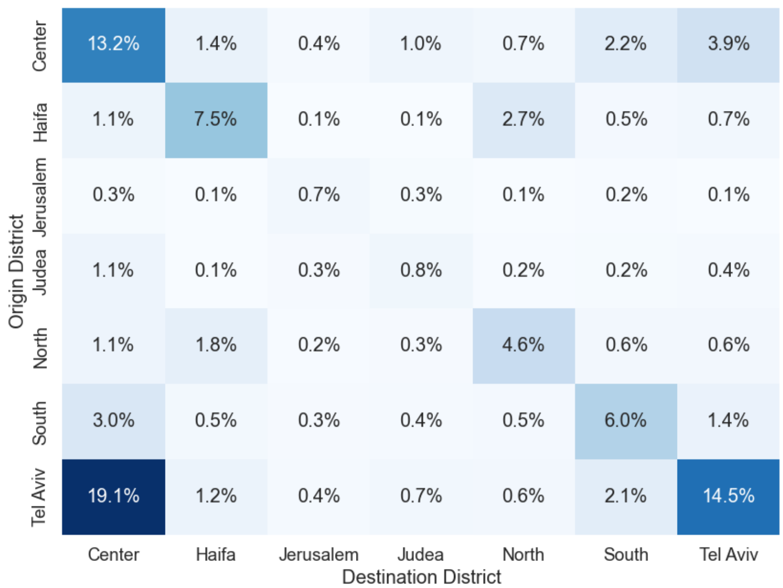


(b) Migrants

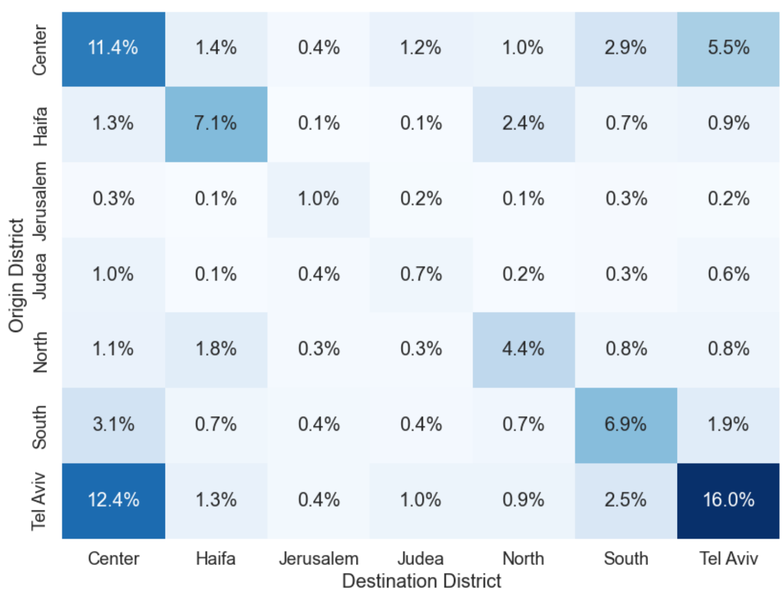
Note: This figure presents the flows of families between localities, for the whole sample of movers, both one and two-time movers. Panel (a) presents the moves of native-born children, while Panel (b) presents the data for migrants. Each plot visually captures the flows from the Former Soviet Union and between the different districts within Israel, delineated into 7 aggregated districts. The diagram is organized from left to right, illustrating the initial move from location A to location B and, for those applicable, a subsequent move from location B to location C. The size of each bar is proportional to the number of movers originating from or destined for each district. Grey lines that connect these bars indicate the flows between districts; the width of each line corresponds to the number of individuals making that specific transition.

Figure C.2: First move transition matrices of native-born, comparing one and two time movers

(a) First move of one-time movers



(b) First move of two-time movers



Note: These tables present the transition probabilities between different locations, delineated into 7 aggregated districts within Israel for native-born children. Rows represent the origin and columns represent the destination districts. Each cell contains the percentage of children making a specific transition, with all cell values in a given matrix summing to 100 percent. Panel (a) presents the first move for native-born children who moved once during childhood, while Panel (b) focuses on the first move made by native-born children who moved twice.

Table C.3: Across-city heterogeneity in childhood location effects on income rank at age 28, single move

	All cities			Overlap cities			
	Cities	Mean	Std.	Cities	Mean	Std.	χ^2 test $H_0 : \theta_j = \theta_1 \forall j$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(i) By θ and η							
Locals							
Cons.	142	0.154	0.214	92	0.180	0.215	143.9
		(0.131)	(0.065)		(0.123)	(0.053)	[0.0004]
Rank-parents	142	-0.004	0.003	92	-0.003	0.003	150.7
		(0.001)	(0.001)		(0.001)	(0.001)	[0.0001]
Immigrants							
Cons.	93	0.640	0.259	92	0.650	0.260	203.8
		(0.078)	(0.044)		(0.047)	(0.044)	[0.0000]
Rank-parents	93	-0.005	0.003	92	-0.005	0.003	204.2
		(0.001)	(0.000)		(0.001)	(0.000)	[0.0000]
(ii) Total city effect							
Locals							
P_{25}	142	0.056	0.178	92	0.1274	0.1768	134.4
		(0.115)	(0.055)		(0.1109)	(0.0465)	[0.0026]
P_{75}	142	-0.138	0.186	92	-0.0153	0.1806	131.2
		(0.112)	(0.055)		(0.1107)	(0.0417)	[0.0046]
Immigrants							
P_{25}	93	0.521	0.235	92	0.5977	0.2358	193.1
		(0.065)	(0.042)		(0.0329)	(0.0421)	[0.0000]
P_{75}	93	0.284	0.273	92	0.4053	0.2743	187.0
		(0.066)	(0.042)		(0.0388)	(0.0423)	[0.0000]

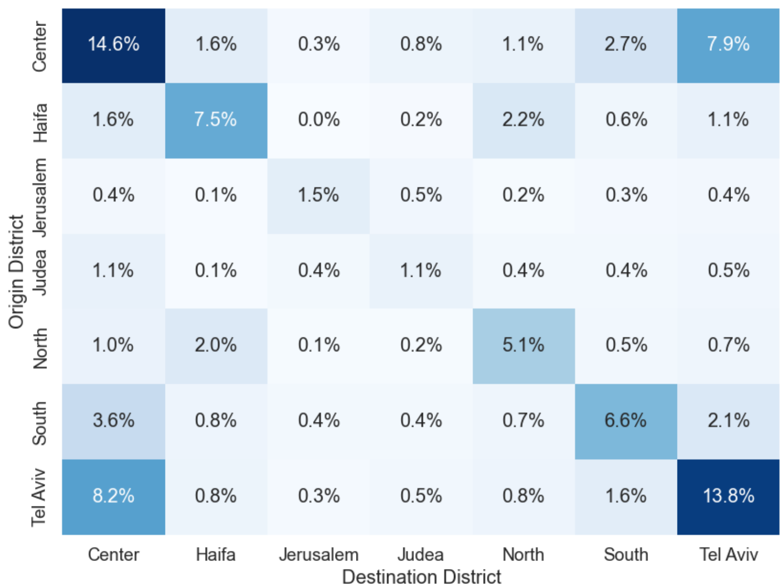
Note: This table presents the estimated bias-corrected standard deviation of the location effect of immigrants and locals. Columns 1-3 present the counts, means, and standard deviation of all the cities that had at least 100 children, and columns 4-6 present the counts, means, and standard deviation for the set of 92 overlapping cities for which we have estimates for both immigrants and locals. Panel (i) displays the θ_{jg} and η_{jg} estimates, and panel (ii) displays the total location effect for families in the 25th and 75th percentile. Column 7 presents a χ^2 test statistic and associated p-value of the null of no location effect heterogeneity across cities. Standard errors for all variance estimators are based on the asymptotic variance, assuming the location effects are drawn from a normal distribution.

Table C.4: Within city heterogeneity between immigrants' and locals' location effects, single move

	Covariance	Correlation	Implied OLS coefficient	Difference		
				Mean	Std.	χ^2 test $H_0: \theta_{jI} - \theta_{jL} = c \; \forall j$
	(1)	(2)	(3)	(4)	(5)	(6)
θ	-0.0047	-0.0839	-0.069	0.5189	0.3571	156.6
	(0.0180)	(0.3213)	(0.266)	(0.1346)	(0.0797)	[0.0000]
		[-0.6407, 0.4661]	[-0.505, 0.335]			
η	0.0000	0.6399	0.542	-0.0013	0.0025	118.3
	(0.0000)	(0.2040)	(0.187)	(0.0015)	(0.0010)	[0.0338]
		[0.2605, 0.9901]	[0.235, 0.923]			
P_{25}	-0.0105	-0.2528	-0.190	0.486	0.340	159.3
	(0.0143)	(0.3396)	(0.256)	(0.118)	(0.072)	[0.0000]
		[-0.845, 0.320]	[-0.643, 0.180]			
P_{75}	0.0014	0.0286	0.019	0.420	0.341	142.6
	(0.0142)	(0.2862)	(0.189)	(0.119)	(0.075)	[0.0006]
		[-0.511, 0.575]	[-0.322, 0.356]			

Note: This table reports the relationship between the location effects of immigrants and the location effects of locals and the test for within-city heterogeneity. Column (1) presents the covariance estimate, column (2) presents the bias-corrected correlation, which is the covariance divided by the standard deviation of immigrants times the standard deviation of locals, and column (3) presents the implied OLS coefficient, which is the covariance divided by the variance of immigrants. Column (4) presents the mean within-city gap between immigrants and locals, column (5) presents the standard deviation of the within-city gap, and column (6) presents the chi^2 test statistic Nd associated p-value of the null of no demeaned within-city gap heterogeneity. Standard errors of the variance and covariances are based on the asymptotic variance, assuming location effects are drawn from a normal distribution. Standard errors of the correlations and OLS slopes are calculated using the delta method. Squared brackets display parametric bootstrapped equal-tailed confidence intervals.

Figure C.3: Second move transition matrix of native-born



Note: This table presents the transition probabilities between different locations, delineated into 7 aggregated districts within Israel for native-born children in their second move during childhood. Rows represent the origin, and columns represent the destination districts. Each cell contains the percentage of children making a specific transition, with all cell values in the matrix summing to 100 percent.

D Research Design Validation

The credibility of our approach depends on the functional form assumptions regarding the relationship between location effects and exposure time and the identification assumption that allows us to identify location effects by estimating Equation 3. In this section, we provide a series of specification tests aimed at validating our research design and supporting our identification strategy.

D.1 Balance Test

The identification assumption requires that exposure time and move age are not systematically correlated with time-invariant factors, such as ability, or time-varying factors, such as parents' investments, that affect the child's income in adulthood. Figures D.1, D.2, D.4, and D.3 provide our first test for these assumptions.

Figure D.1 presents the relationship between native-born children's age when the family moved for the first time and parents' characteristics. In sub-figures D.1b and D.1a, the blue dots and error bars represent the raw relationship between the age at the time of the move and parents' education for native-born Israeli children and the gray ones present the same relationship after controlling for the set of covariates mentioned in Section 3.2. Parents' years of schooling are obtained from the 1995 census, which is available for 20% of the population. Consistent with findings from Heckman and Landersø (2021), we find that the more educated the parents, the more likely they are to move when their children are younger. However, after controlling for x_i, we find that this relationship disappears, supporting our identification assumption of no systematic relationship with time-invariant characteristics.

Sub-figure D.1c displays the relationship between the age of move and parents' earnings growth when the child was between ages 0-5. To measure this variable, we use an earnings dataset that is available only for salary-employed workers in the years 1986-1995. Thus, since we don't use this dataset to construct our parental income, we avoid mechanical correlation between parents' rank and wage growth up to age 5.[36] This figure reveals that families that moved when the children were in older ages experienced higher wage growth when the child was between 0-5 years old. However, after controlling for our set of fixed effects and controls,

[36] We don't use the salary-employed earning data for the years 1986-1995 because it does not include information on earnings from self-employment.

we find that there is no statistically significant relationship between the age of move and parents' earnings growth at younger ages. This suggests that time-varying components are also not systematically correlated with the child's age at the move after adjusting for location choices and parents' income rank.

Next, in Figure D.2, we examine the relationship between the child's age at arrival to Israel and the parents' year of schooling. In sub-figures D.2b and D.2a, the sample is restricted to all the immigrants who arrived in Israel before 1995 and answered the demographic survey in the 1995 census. Similar to the natives, in the general population of immigrants, a child's age of arrival is negatively correlated with parents' years of schooling, but this relationship disappears after adding controls.[37]

In Figures D.4 and D.3, we repeat the exercise, now assessing whether the age of the child at the second move is related to the parents' characteristics. Interestingly, we find that among the families that moved twice, there is a very weak relationship between the child's age at the second move and the parents' education also in the uncontrolled model.

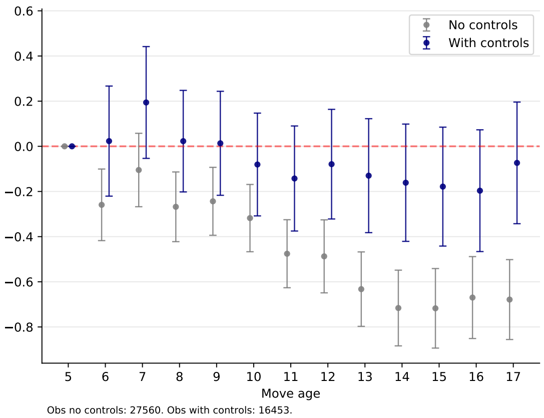
D.2 Uncovering the Functional Form

Moving to cities with higher mean outcomes - To test whether the effects of childhood location follow a linear relationship as modeled in Equation 1, and to identify the age A at which childhood location no longer impacts children's outcomes, we proceed in two steps. In the first step, following the benchmark diagnostics in the literature (Chetty and Hendren, 2018a), we leverage a split sample methodology. We test whether relocating at an earlier age to a city with higher city-level mean permanent residents outcomes increases the child's long-run outcomes. To minimize the impact of measurement errors that could attenuate the results, we sidestep heterogeneity related to parents' rank.

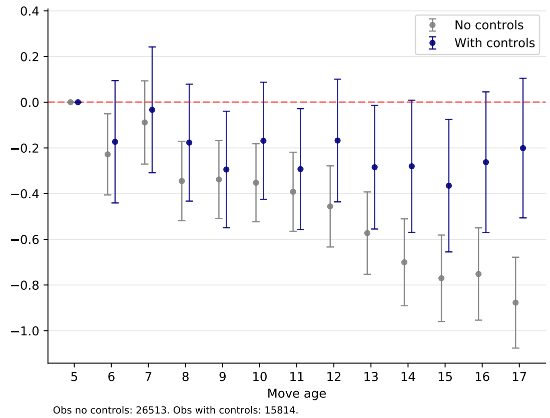
For the non-immigrants, we estimate for every city j the city-level mean income rank at age 28 $\bar{Y}_j$ among the children whose families did not move between cities until the age of 30. Then, in our main analysis sample of natives who moved only once between cities in childhood, we study whether moving one year earlier from origin o to destination d shifts outcomes in the same direction as the difference between origin and destination city mean outcomes and how this relationship varies by the age of the child. Formally, on the sample

[37]We do not present the wage growth balance because, for most of the immigrant families, we do not observe parental wages when the children are as young as 5 years old.

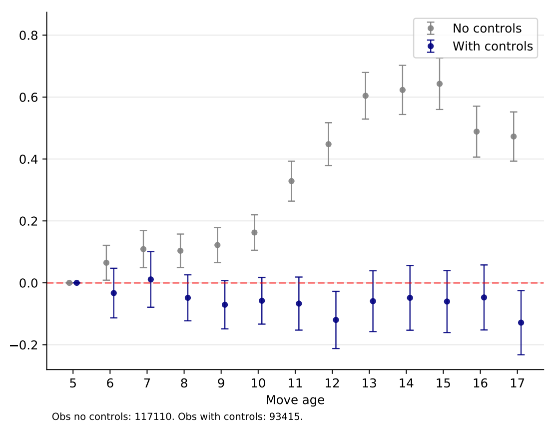
Figure D.1: Relationship between age at first move and parents' characteristics, native-born children



(a) Mother years of schooling



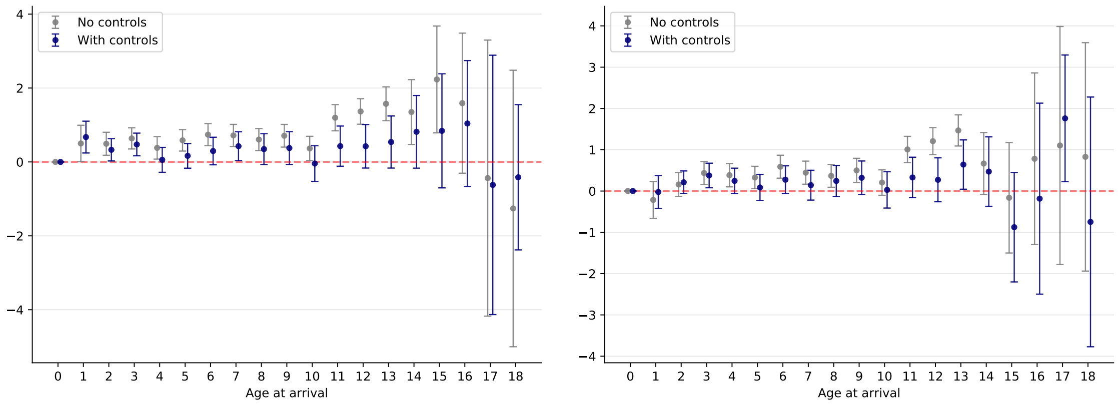
(b) Father years of schooling



(c) Wage growth, ages 0-5

Note: This figure presents balance test results for the non-immigrant sample, showing the relationship between the child's age at the first move between cities in Israel and parents' characteristics. Sub-figures D.1a and D.1b present parents' years of schooling, as recorded in the 1995 census. Sub-figure D.1c presents the relationship between the age of move and parents' earnings growth when the child was aged 0-5. Controls include origin-destination fixed effects, birth-year fixed effects, and parents' income rank interacted with the year of birth. Confidence intervals are constructed using family levels clustered standard errors.

Figure D.2: Relationship between age of arrival to Israel and parents' characteristics, immigrants



(a) Father years of schooling

(b) Mother years of schooling

Note: This figure presents the relationship between the child's age at the move to Israel and the parents' education. Sub-figure D.2a presents the father's years of schooling, and D.2b presents the mother's years of schooling from the 1995 census. Controls include origin-destination-birth-year fixed effects and parents' income rank interacted with the year of birth. Confidence intervals are constructed using family levels clustered standard errors.

of native-born children to families who moved once, we run the following regression:

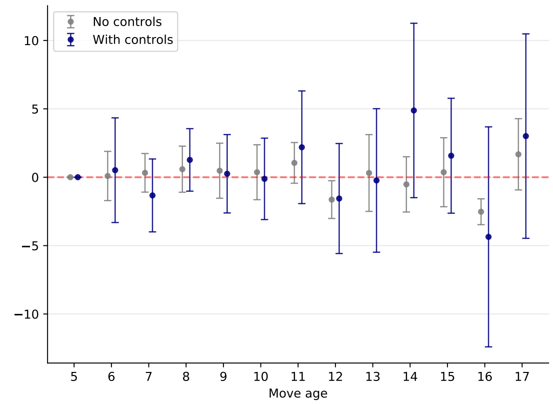
$$Y_i = \sum_{m=1}^{30} \beta_m \mathbb{1}\{m(i) = m\} \Delta_{o(i)d(i)} + x_i' \gamma + \epsilon_i, \tag{19}$$

where $o(i)$ is the origin city of child i, $d(i)$ is the destination city of child i, x_i includes all the controls and fixed effects described in section 3.2 which are origin-destination-second destination fixed effects, and cohorts dummies interacted with parents' income rank, and $\Delta_{od} = \bar{Y}_d - \bar{Y}_o$. Our parameters of interest are β_m coefficients, which, under our identification assumption, measure the effect of moving at age a to a destination city d, which has one percentile rank higher children income rank in age 28 than in the origin city o.

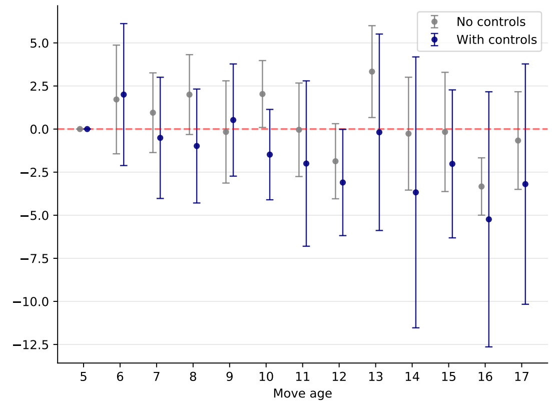
For immigrants, we conduct a similar exercise, with the adjustment for the fact that immigrants include both children who were born in Israel and children who arrived in Israel from the USSR at different ages. We calculate $\bar{Y}_j$ among the children of immigrants who were born in Israel and stayed in the same city up until age 17. Because our main analysis sample includes all immigrants, including those who were born in Israel,[38] we randomly split the

[38]They are included with a moving age of zero, that is, with exposure of 17 years to their location.

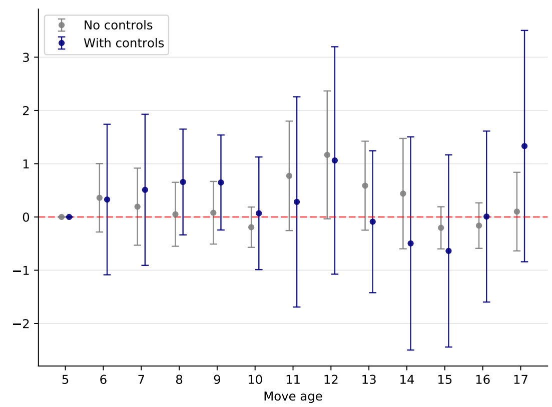
Figure D.3: Relationship between age at second move and parents' characteristics, native-born children



(a) Mother years of schooling



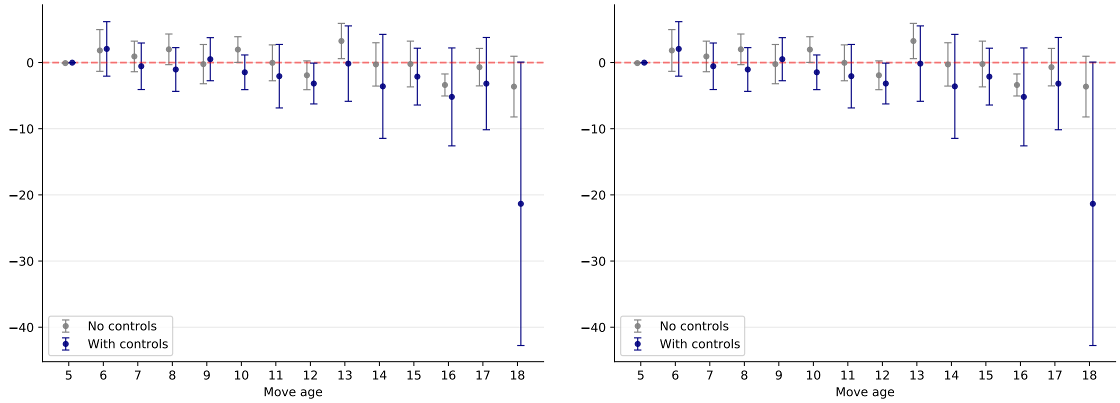
(b) Father years of schooling



(c) Wage growth ages 1-5

Note: This figure presents balance test results for the non-immigrant sample, showing the relationship between the child's age at the second move between cities in Israel and parents' characteristics. Sub-figures D.3a and D.3b present parents' years of schooling, as recorded in the 1995 census. Sub-figure D.3c presents the relationship between the age at the second move and parents' earnings growth when the child was aged 0-5. Controls include origin-first destination-second destination fixed effects, birth-year fixed effects, and parents' income rank interacted with the year of birth. Confidence intervals are constructed using family levels clustered standard errors.

Figure D.4: Relationship between age when the family moved between cities in Israel and parents' characteristics, immigrants



(a) Father years of schooling

(b) Mother years of schooling

Note: This figure presents the relationship between the child's age at the move to Israel and the parents' education. Sub-figure D.4a presents the father's years of schooling, and D.4b presents the mother's years of schooling from the 1995 census. Controls include origin-destination-birth-year fixed effects and parents' income rank interacted with the year of birth. Confidence intervals are constructed using family levels clustered standard errors.

sample of immigrants who were born in Israel into two random samples, and within each sample $s \in \{1, 2\}$, we calculate their mean outcome $\bar{Y}_{js}$. Then, on the sample of immigrant children who stayed in the same city until age 18, we run the following regression:

$$Y_i = \beta_0 \mathbb{1}\{a(i) = 0\}\bar{Y}_{j(i)s'(i)} + \sum_{a=1}^{17} \beta_a \mathbb{1}\{a(i) = a\}\bar{Y}_{j(i)} + x_i'\gamma + \epsilon_i, \tag{20}$$

where $s'(i)$ is the random sample that does not include child i. Our parameters of interest are the β_a coefficients, which measure the effect of arriving in Israel at age a to a city with one percentile rank higher average income of immigrants born in Israel. Note that our cohort restriction, together with the fact that the immigration wave is restricted to the years 1989-2000, implies that we have only a few observations with an immigration age above 17.[39]

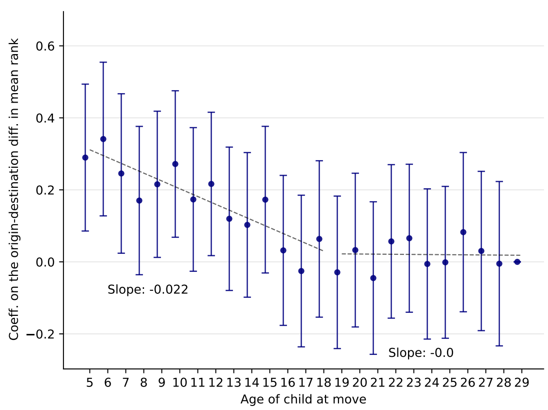
Figure D.5 plots the estimates and confidence intervals of β_m from both Equations 19 and 20, with the earnings rank at age 28 as the outcome.

The estimates in Figure 19 on the non-immigrants exhibit two key patterns: First, the estimates of β_m decline steadily with ages $m < 18$ at a linear rate, suggesting that the age

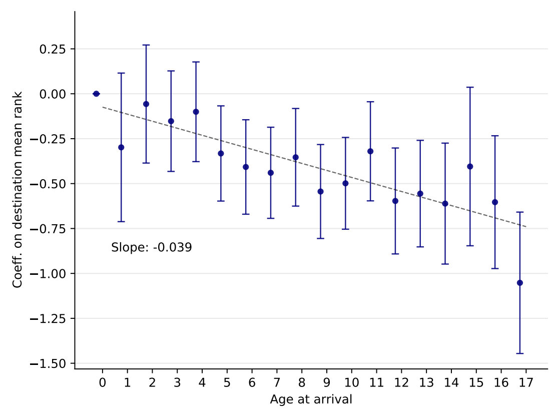
[39] See Appendix figure A.3 for immigrants' age at arrival distribution.

of arrival to a city with higher mean outcomes is approximately linearly related to children's income rank at age 28. Such a pattern was also found in the US (Chetty and Hendren, 2018a), Australia (Deutscher, 2020), and Canada (Laliberté, 2021). Second, we find that for ages $m \geq 18$, the effect of parents moving to a location with one percentage higher outcome of permanent residents in the destination compared to origin does not change with the age of the child when moving. Such a result is in line with the institutional setting in Israel in which most of the children enlist in the Israeli army at ages 18-19 and, therefore, are significantly less exposed to the local institutions and peers.

Figure D.5: Childhood exposure effects on earnings rank in adulthood



(a) Natives



(b) Immigrants

Note: This figure presents the exposure effect coefficients on earnings rank measured at age 28, against the child's age at the time of move. In dashed lines are the linear piece-wise fitted lines with the corresponding slope. Panel (a) presents the effects on native-born Israeli children, estimated via Equation 19, and Panel (b) presents the effects on immigrants from the former Soviet Union, estimated via Equation 20. Earning ranks are defined as percentiles within cohort. Controls include origin-destination-birth-year fixed effects and parental income rank interacted with the year of birth. Confidence intervals are constructed based on family-level clustered standard errors.

Similarly, the pattern in Figure D.5 among immigrants displays a similar pattern where b_a declines steadily with age. Regressing $\hat{b}_m$ of m for $m \leq 18$ among native-born children, we estimate a slope of 0.022. That is, the income rank at age 28 of children who moved gets similar to the mean income rank of permanent residents at a rate of 2.2% per year until age 18. Repeating the same exercise for immigrants, we estimate a larger slope of 0.039. There could be several explanations for why the exposure slope among natives is larger than the exposure slope among immigrants. Following our findings above, we find that the location effects of immigrants are larger than the location effects of native-born. Therefore, it is likely that our exposure-sloped results express that to some degree. Second, it is also possible that sampling error in $\bar{Y}_j$ attenuates our estimated slope more severely for natives since their sample is smaller, but they also have more controls.[40]

The kink at age 18 motivates a piece-wise regression to estimate the slope coefficients on the individual-level data formally. For natives, we use the following specifications:

$$\begin{aligned} Y_i =&(\gamma_{below} + \beta_{below} m(i))\Delta_{o(i)d(i)}\mathbb{1}\{m(i) \leq 18\} \\ &+ (\gamma_{above} + \beta_{above}(30 - m(i)))\Delta_{o(i)d(i)}\mathbb{1}\{m(i) > 18\} + x_i'\gamma + \epsilon_i, \end{aligned} \tag{21}$$

where the variable definitions and controls are identical to the ones used in Equation 19. The coefficients of interest are β_{below} for the slope below age 18 and β_{above} for the slope above age 18, which we expect to be zero. The corresponding regression equation for immigrants is the following, where we only estimate the slope up to age 18, due to our data constraints:

$$Y_i =\gamma_{below} + \beta_{below} m(i)\bar{Y}_{j(i)s'(i)} + x_i'\gamma + \epsilon_i, \tag{22}$$

where definitions follow Equation 20.

Table D.1 presents the estimated slopes for natives and immigrants in columns (1) and (3). First, we find a small and non-significant coefficient for the slope above age 18 for locals. Additionally, we find a slope of 0.022 and 0.033 for natives and immigrants up to age 18, respectively. These estimates are comparable yet smaller than the effects found in previous studies conducted in the US (0.035 at the county level in Chetty and Hendren (2018a)), Australia (0.033 in Deutscher (2020)), and Canada (0.042 in Laliberté (2021)).

Robustness to family fixed effect - Our key identifying assumption is that the potential outcomes of children who move to better vs. worse cities do not vary with the age of move or the age of immigration. If families with higher mean ability move to cities when the

[40]As depicted in the descriptive statistics Table 1, the sample of immigrants is larger than the sample of native-born children who moved in childhood. In contrast, the number of controls in the natives' regression is larger because there are more origin-destination fixed effects in the natives' sample than in the immigrants' sample since the origin location of all the immigrants is always the USSR. Therefore, it is possible that our coefficients are more attenuated due to the sampling error in $\bar{Y}_j$. One could correct the sampling error using an error in variable regression or split sample 2SLS. However, due to computational restrictions in the computer and statistical software provided by the Israeli Central Bureau of Statistics, we could not implement them. Future drafts will address these issues.

Table D.1: Relationship between years of exposure in childhood city and mean outcomes and posterior location effects

	Mean outcomes				Posteriors	
	Natives (1)	Natives (2)	Immigrants (3)	Immigrants (4)	Immigrants (5)	Natives (6)
$\Delta\times$ below 18	-0.022	-0.019	-0.033	-0.027	-1.433	-0.743
	(0.005)	(0.011)	(0.006)	(0.013)	(0.295)	(0.204)
$\Delta\times$ above 18	0.006	0.005				
	(0.006)	(0.008)				
Family fixed effect	No	Yes	No	Yes	No	No
Obs.	95,500	70,549	138,664	110,462	138,664	95,500

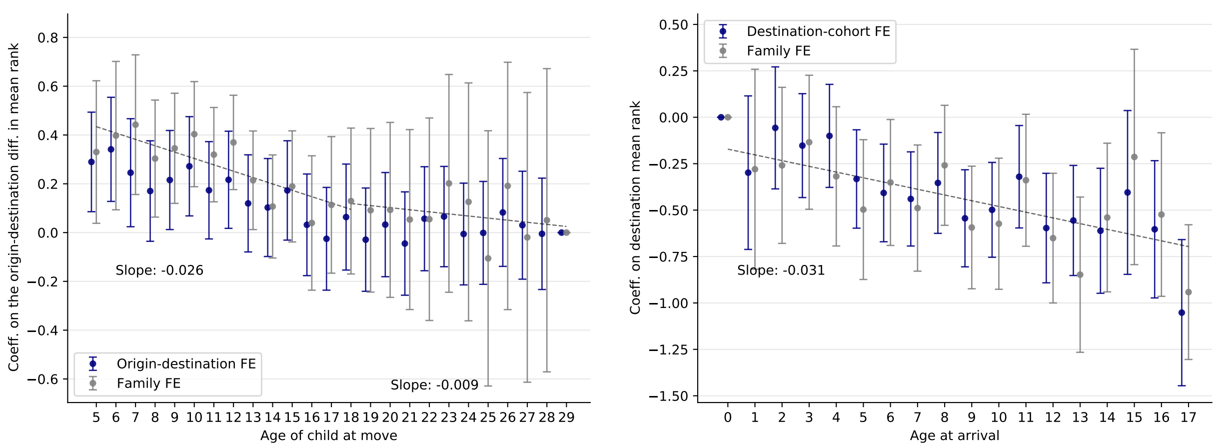
Note: This table reports the linear slope coefficients between location effects or posterior means and the age of move. Columns (1) and (3) present the effects estimated via Equations 21 and 22, accordingly. In columns (2) and (4) we add family fixed-effects to the estimation. Immigrants columns do not have coefficients for moves above age 18 due to our sample restriction. Columns (5) and (6) present the slope coefficients between the posterior means and the arrival age for immigrants and the age of move for locals. Standard errors in parenthesis are clustered at the family level.

children are younger, then Assumption A1 is violated. To test this, we control for differences in family-level factors by including family fixed-effects when estimating Equations (19) and (20). The inclusion of family fixed effects implies sibling comparisons.

Figure D.6 displays the estimates from Equations (19) and (20), with and without family fixed effects on child income rank at age 28. The blue dots and confidence intervals replicate the estimates from Figure D.5, without controlling for family fixed effects. The gray dots and confidence intervals present the estimates with family-fixed effects. The linear decline in the estimated values of β_m for locals and β_a for immigrants until age 18 is very similar to that in the baseline specification, although noisier. Siblings who moved to a city with high outcomes at younger ages have better outcomes than their older siblings. Table D.1 presents in columns (2) and (4) the corresponding slopes estimated with family fixed effects. We find coefficients of 0.19 for natives and 0.026 for immigrants, values close to the estimates above without family fixed effects, presented in the same table in columns (1) and (3).

Placebo test using outcomes realized in childhood - In Appendix Figure D.7, we exploit outcomes that are realized before age 18 as a placebo test. If the relationship between child's outcomes and city-level mean outcomes we estimate is driven by selection, that is, that higher ability children move to higher mean outcome cities at younger ages, then we should expect to see the same linear pattern of decreasing impacts until age 18 for every

Figure D.6: Robustness to family fixed effects



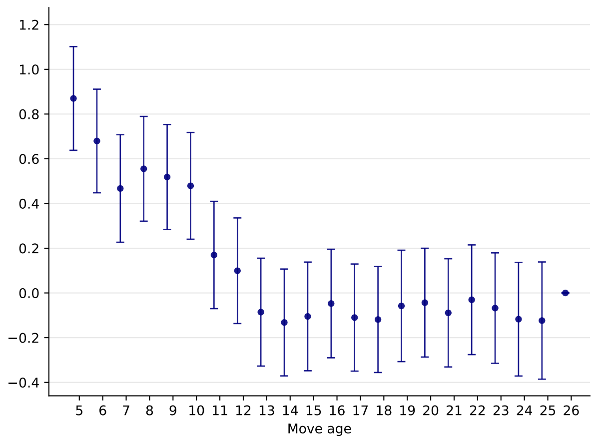
(a) Natives

(b) immigrants

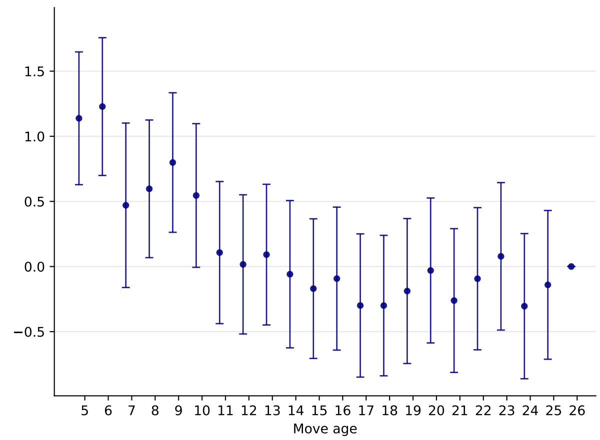
Note: This figure presents the exposure effect coefficients on earnings rank measured at age 28 against the child's age at the time of move. In dashed lines are the linear piece-wise fitted lines with the corresponding slope. Panel (a) presents the effects on native-born Israeli children, estimated via Equation 19, and Panel (b) presents the effects on immigrant children, estimated via Equation 20. In grey, we present the results of a regression with family fixed effects instead of the origin-destination-birth-year fixed effects, which implies sibling comparison. In blue, we present the results of the regressions as presented in Figure D.5 for comparison. Earning ranks are defined as percentiles within year. Confidence intervals are constructed based on family-level clustered standard errors.

outcome, regardless of its realization year. Otherwise, we expect the kink in the effects from which values flatten out to appear sooner, around the age of realization.

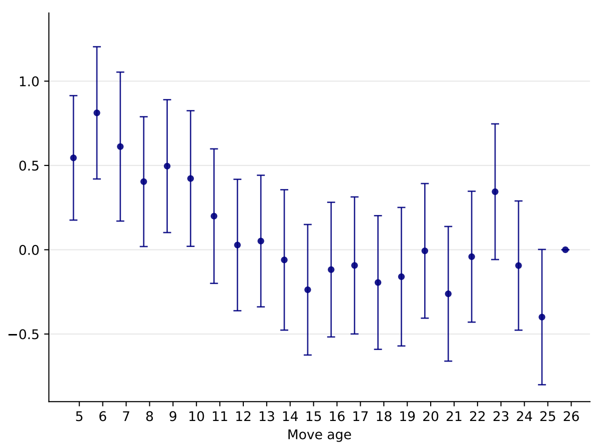
Figure D.7: Childhood exposure effects on 5th grade standardized exam scores of natives



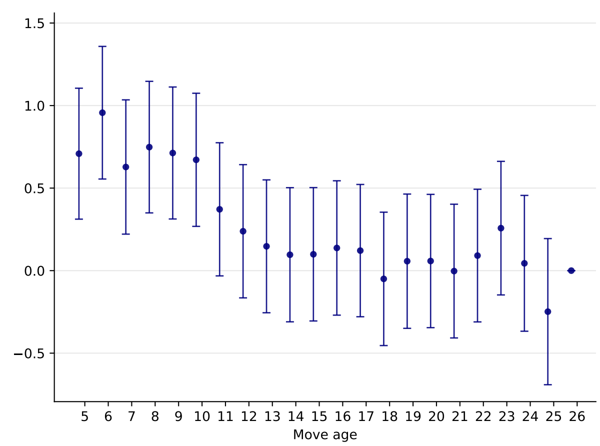
(a) Math



(b) Science



(c) English



(d) Hebrew

Note: This figure presents the exposure effect coefficients on the national standardized exam scores in the 5th grade, the *Meitzav*, which mimics the PISA exam administrated by the OECD. The different panels display the results using students' scores in different key subjects: mathematics (a), science (b), English (c), and Hebrew (d), measured in standard deviations. This exam is administered at a random representative sample of schools, which in our sample amounts to a tenth of the children in the full analysis sample. We present the effects on non-immigrants, estimated via Equation 19. Confidence intervals are constructed based on family-level clustered standard errors.

For this purpose, we run Equation 19 utilizing children's scores in the 5th-grade national standardized school evaluation exams, called *Meitzav*, administered by the Israeli Ministry of Education.[41] The *Meitzav* exams cover several key subjects, including mathematics, science, English, and Hebrew or Arab language skills, and closely mirror the PISA exam administrated by the OECD. This exam is administered at a random representative sample of schools, which in our sample amounts to a tenth of the children in the full analysis sample. Since the exam is taken when the child is at age 11, the estimates in years 11-18 serve as a placebo test as we expect to find null effects in this age span. Note that because children who immigrated after the age of 11 never took the exam, we can perform this exercise only among the non-immigrants. The results in Figure D.7 validate our assumption, as we find that the effect of moving to a city with higher mean grades in the Meitzav exam declines with the age of move up to age 11 and stabilizes thereafter.

Test for linearity - The uncontrolled mean outcomes used in the above regressions consist of both location effects and the mean ability $\bar{Y}_j = \theta_j + \bar{\xi}_j$. That is, the relationship estimated in Figure D.5 could reflect either linearity with respect to location effects or a linear relationship between the age of move and location effects along with a selection bias. Therefore, the interpretation of the estimates in the above exercises is mixed. On the one hand, it does tell us that moving to a place with higher mean outcomes of permanent residents is approximately linear with the years of exposure to the location. However, it is less clear if the location effects θ_{jg} themselves are linearly related to the years of exposure. If $\bar{\xi}_o - \bar{\xi}_d = 0$, then the slope $\beta_{m+1} - \beta_m$ with respect to child's age should be approximately $\frac{1}{A} \approx 0.055$ for $A = 18$. Otherwise, it is the sum of $\frac{1}{A}$ and the relationship between nonmovers' ability and location effects.

To address this concern, we conducted an additional exercise. We regress the estimated location effects of movers from Equation (3), i.e., the estimated location effects residualized of selection based on x_i, on children's long-run outcomes, allowing separate coefficient for each age at the time of the move. If the relationship is approximately linear, we should expect a constant rate of change in the age-specific coefficients by the age at move.

Since Equation (3) is estimated on the same sample, we expect a mechanical correlation. To avoid this, we run Equation (3) on two random splits $s(i) \in \{1, 2\}$ of our analysis sample. Given our findings above, we run these regressions using age $A = 18$ as the last age at which places affect outcomes. Therefore for every group g, city i, and sample split s, we estimate

[41] "Meitzav" is a Hebrew acronym translating to "School Efficiency and Growth Measures".

$\hat{\theta}^{s}_{jgp}$. Then, to test for the linearity assumption on the sample of natives, we run:

$$Y_i = \sum_{m=1}^{18} \beta_m \mathbb{1}\{m(i) = m\} \hat{\Delta}^{s'(i)}_{o(i)d(i)p(i)} + x_i'\gamma + \epsilon_i \tag{23}$$

where $s'(i)$ is the random sample that does not include child i, $\hat{\Delta}^{s}_{od} = \hat{\theta}^{*s}_{dgp} - \hat{\theta}^{*s}_{ogp}$, and $\hat{\theta}^{*s}_{dgp}$ are the Empirical Bayes (EB) posterior mean of θ_{dgp} we estimate in Section 8. Similarly, we estimate the equivalent regression on the sample of immigrants:

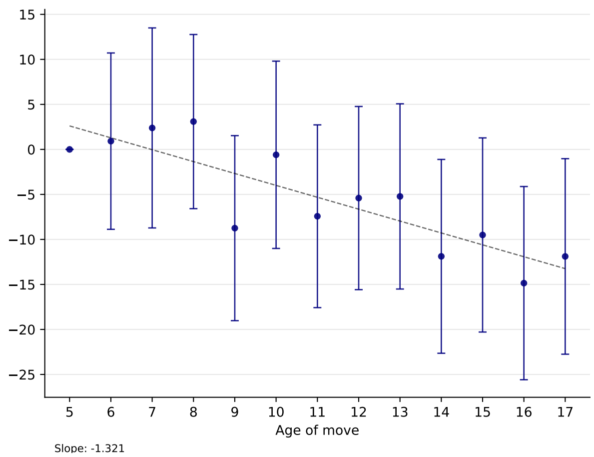
$$Y_i = \sum_{a=1}^{18} \beta_a \mathbb{1}\{a(i) = a\} \hat{\theta}^{*s}_{j(i)g(i)} + x_i'\gamma + \epsilon_i \tag{24}$$

where the base-level age is the immigrants who were born in Israel.

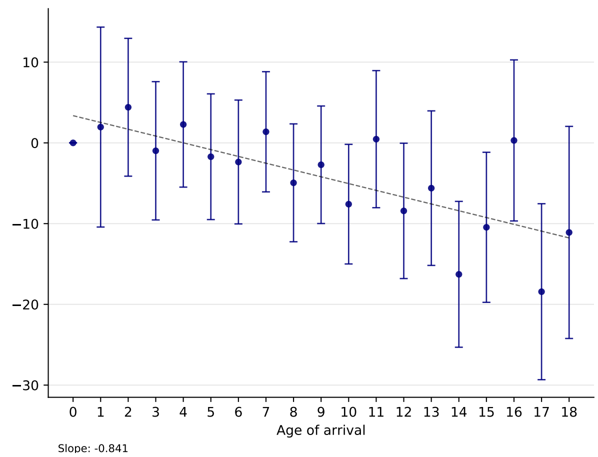
If Equation 3 captures the true functional form of location effects and outcomes, then for immigrants, the coefficient β_a should give us the *number* of θ_j accumulated in location j each year of exposure as θ^* is the effect of spending one year in city j. Thereby, the slope $\beta_{a+1} - \beta_a$, with respect to exposure time should be 1 on average. Accordingly, for natives, β_m gives the equivalent sum of location effects of the destination location relative to the origin.

Figure D.8 presents the main results visually. We observe a linear relationship between the posterior means and the age of move for both native-born and immigrants, with fitted slopes of 1.321 and 0.841, respectively. In Table D.1, we present the formally estimated slopes in column (5) for immigrants and column (6) for locals, with values of 1.433 and 0.743, respectively. Importantly, in support of the linear effect assumption, we cannot reject the null that these coefficients are different than 1.

Figure D.8: Relationship between age of arrival/move and posterior mean



(a) Native-born



(b) Immigrants

Note: This figure displays the relationship between the years of exposure and the posterior mean of location effect estimated via Equations (23) for immigrants and (24) for native-born Israeli children. We present the β_a coefficients for each arrival age for immigrants in Panel (a) and β_m coefficients for each moving age in Panel (b), with the confidence intervals based on robust standard errors as vertical lines. The dashed lines represent the linear fitted lines, and their slopes are at the bottom left of the figure. The baseline coefficient is the lowest age of arrival/move.

E Robustness Exercises

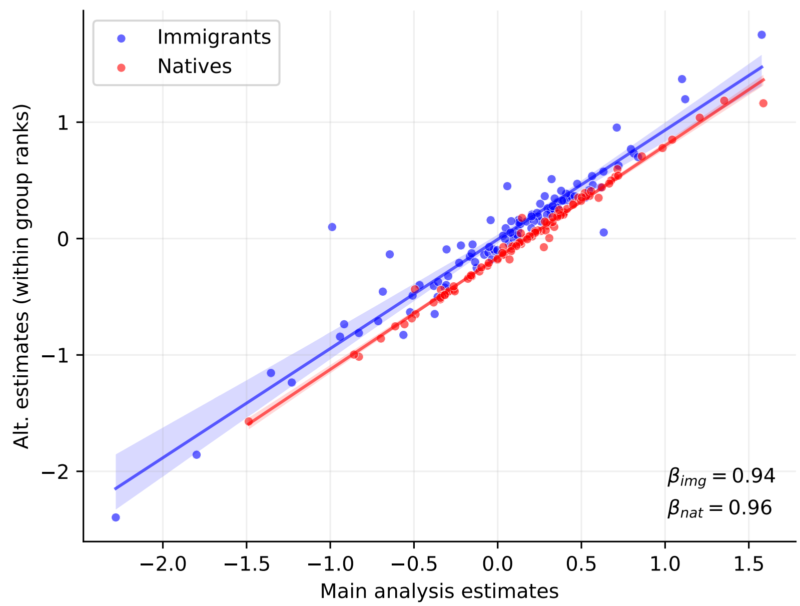
E.1 Within Group Parental Income Rank

Immigrants face earnings penalties due to frictions such as language barriers, cultural differences, and lack of network and information. Therefore, if the heterogeneity in location effects with respect to parental income is driven by heterogeneity in skills, we classify high-earning potential parents as low earnings. By doing so, when we compare the location effects of immigrants and natives, we do not compare families with the same set of skills. To test the sensitivity of our results to the mismeasurement of parental income of immigrants, we run our main model in Equation 3, but instead of measuring parental income rank at the national distribution, we measure it separately within group.

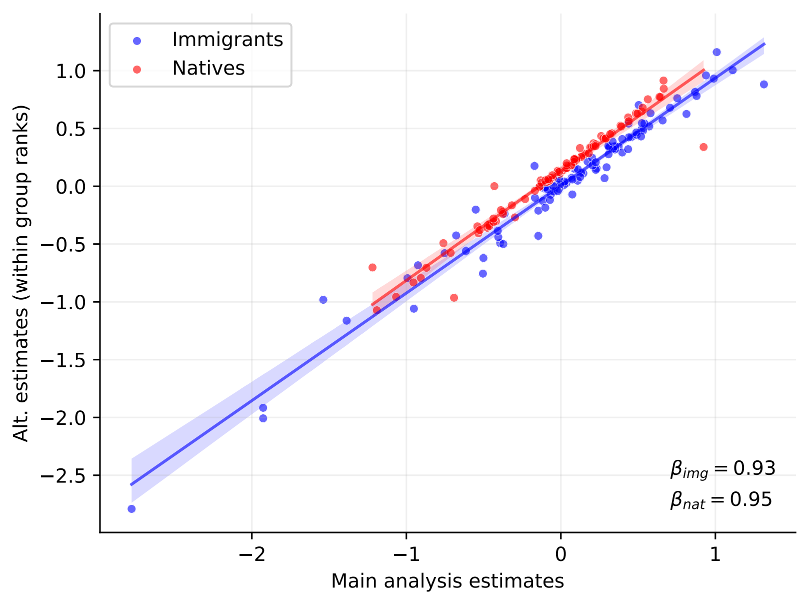
Table E.1 presents the resulting correlation matrix of immigrants' and natives' location effects under that definition of parental income rank. The correlation structure is qualitatively similar, with a correlation between immigrants and natives of -0.12, which is lower by four percentage points, suggesting that the negative relationship between immigrants and natives is not driven mechanically by classifying immigrants with low-income relative to natives but high-income relative to other immigrants as low-income families.

Figure E.1 presents the scatter plots and regression lines of the alternative within-group parental ranks against our main analysis estimates based on ranks in the population. The two panels plot the estimates for low and high-income families. We find that the correlation between the estimates is strong, with slopes of well above 0.9 for both immigration groups.

Figure E.1: Location effects using within-group parental ranks instead of population ranks



(a) 25th percentile



(b) 75th percentile

Note: This figure presents the relation between location effects estimated based on parental ranks defined either with respect to the full population distribution (portrayed on the X-axis, our main analysis definition), or with ranks defined within the group distributions, immigrants, or natives (depicted on the Y-axis). Panel (A) presents the estimates for low-income families, while Panel (B) presents the corresponding effects for high-income families. In the bottom corner are the slope coefficients of the linearly fitted lines.

Table E.1: Correlation matrix of location effects of immigrants and natives, within group parents income rank

	θ_{25}		θ_{75}	
	Natives (1)	Immigrants (2)	Natives	Immigrants
Natives	0.158 (0.052)		0.193 (0.041)	
Immigrants	-0.116 (0.338) [-0.689, 0.439]	0.186 (0.026)	0.488 (0.215) [0.064, 0.882]	0.248 (0.025)

Note: This table reports the standard deviation in diagonal and correlation in off-diagonal of immigrants' and natives' location effects on children's income rank at age 28. Columns 1-2 display the correlation matrix for low-income families from the 25th percentile of the within-group income distribution, and columns 3-4 display the correlation matrix for high-income families at the 75th percentile of the within-group income distribution. Standard errors of the variance and covariances are based on the asymptotic variance, assuming location effects are drawn from a normal distribution. Standard errors of the correlations and standard deviations are calculated using the delta method. Square brackets display parametric bootstrapped equal-tailed confidence intervals.

E.2 Other Concave Transformation of Income

Table E.2 displays the correlation matrix of location effects on other earnings variables. In panel (A), we run Equation 3 on earnings at age 28 and in panel (B) on log earnings. Panel (A) suggests that the hyperparameter estimates on earnings entail excessive noise. Nevertheless, point estimates align qualitatively with our main results using income rank at age 28. However, when using log earnings (excluding zeros), which, similar to rank, is a concave transformation of earnings, results are in accord with Tables 2 and 3. First, the variance of immigrant location effects is twice the size of the variance of native-born children's location effects. Second, the correlation between immigrants' and natives' location effects is 0.02 for low-income families and 0.56 for high-income families.

Table E.2: Correlation matrix of location effects for immigrants and natives on earnings and log earnings

	P_{25}		P_{75}	
	Natives (1)	Immigrants (2)	Natives	Immigrants
A) Earnings				
Natives	169.53 (300.71)		262.42 380.80)	
Immigrants	-0.75 (1.48) [-3.12, 0.96]	367.35 (61.53)	0.35 (1.18) [-0.81, 1.16]	557.07 (59.98)
B) Log earnings				
Natives	0.0181 (0.0081)		0.0179 (0.0068)	
Immigrants	0.0284 (0.4008) [-0.614, 0.324]	0.0268 (0.0042)	0.5638 (0.3220) [0.122, 1.138]	0.0350 (0.0043)

Note: This table reports the standard deviation in diagonal and correlation in off-diagonal of immigrants' and natives' location effects on children's earnings at age 28 measured in Shekels (Panel A, 1 US $ ≈ 3.4 ILS). Columns 1-2 display the correlation matrix for low-income families from the 25th percentile of the within-group income distribution, and columns 3-4 display the correlation matrix for high-income families at the 75th percentile of the within-group income distribution. Standard errors of the variance and covariances are based on the asymptotic variance, assuming location effects are drawn from a normal distribution. Standard errors of the correlations and standard deviations are calculated using the delta method. Square brackets display parametric bootstrapped equal-tailed confidence intervals.

E.3 Robustness to City Weights

In Table 3, the variance components are calculated after reweighting each city by the total number of residents. To assess the sensitivity of location effect heterogeneity to the weighting scheme, Table E.3 reports the standard deviation and correlations of location effects under alternative weights. Panel (i) presents results when cities are reweighted by the total number of movers in each city, i.e., by the total number of families who are included in our main regression sample. Panels (ii) and (iii) allowed for different weights for immigrants and natives, where panel (ii) reweights by group number of movers and panel (iii) reweights by group number of residents, including both movers and non-movers.

Table E.3: Correlation matrix of location effects of immigrants and natives, robustness to city level weights

	θ_{25}		θ_{75}	
	Natives (1)	Immigrants (2)	Natives	Immigrants
(i) Total # of movers weights				
Natives	0.163		0.171	
	(0.048)		(0.042)	
Immigrants	-0.067	0.264	0.429	0.339
	(0.253)	(0.043)	(0.201)	(0.042)
	[-0.532, 0.383]		[0.070, 0.825]	
(ii) Group # of movers weights				
Natives	0.173		0.198	
	(0.050)		(0.041)	
Immigrants	-0.081	0.205	0.468	0.267
	(0.306)	(0.026)	(0.217)	(0.026)
	[-0.622, 0.457]		[0.025, 0.879]	
(iii) Group # of residents weights				
Natives	0.183		0.163	
	(0.045)		(0.042)	
Immigrants	-0.180	0.190	0.470	0.248
	(0.299)	(0.023)	(0.268)	(0.023)
	[-0.732, 0.360]		[-0.005, 0.970]	

Note: This table reports the standard deviation in diagonal and correlation in off-diagonal of the location effects of immigrants and natives on children's income rank at age 28 for different reweighting schemes. Columns 1-2 display the correlation matrix for low-income families from the 25th percentile of the within-group income distribution, and columns 3-4 display the correlation matrix for high-income families at the 75th percentile of the within-group income distribution. In panel (i), cities are reweighted by the total number of movers; in panel (ii), cities are reweighted by the number of each group's movers; and in panel (iii), cities are reweighted by each group's total number of residents. Standard errors of the variance and covariances are based on the asymptotic variance, assuming location effects are drawn from a normal distribution. Standard errors of the correlations and standard deviations are calculated using the delta method. Square brackets display parametric bootstrapped equal-tailed confidence intervals.

F Estimation of Variance Componenets

For every city $j \in \{1, .., J\}$ in Israel, we denote θ_{jgp} the effect of spending one more year in the city j for an individual who belongs to group $g \in \{I, L\}$, with parent income rank p. We model the location effects in each city j to be:

$$\theta_{jgp} = \alpha_{jg} + \eta_{jg} p,$$

where θ_{jg} measures the returns to the city j for individuals who belong to group g with parents at the lowest income rank, and η_{jg} measures the returns to parents rank in city j. Therefore, for every group g, we end up with a vector of size J of location effects intercepts $\alpha_g = (\alpha_{1g}, ..., \alpha_{Jg})'$, and a vector of size J parents ranks slopes $\eta_g = (\eta_{1g}, ..., \eta_{Jg})'$. Finally, we denote $\theta \in \mathbb{R}^{4J}$ the stacked vector that describes the location effects in Israel $\theta = (\alpha'_{\mathcal{N}}, \eta'_{\mathcal{N}}, \alpha'_{\mathcal{I}}, \eta'_{\mathcal{I}})'$.

We estimate θ by running equation (3) which results with a $4 \cdot J$ vector of estimated location effects $\hat{\theta} = (\hat{\alpha}'_{\mathcal{N}}, \hat{\eta}'_{\mathcal{N}}, \hat{\alpha}'_{\mathcal{I}}, \hat{\eta}'_{\mathcal{I}})'$, and $\Sigma = \mathcal{V}(\hat{\theta})$, the sampling variance of $\hat{\theta}$. In this appendix section, we provide a detailed explanation of how we estimate the variance of θ and its standard error.

F.1 Method of Moments Variance Component Estimate

We denote $\Omega \in \mathbb{R}^{4\times4}$ the variance covariance matrix of $\theta_j = (\alpha_{j\mathcal{N}}, \eta_{j\mathcal{N}}, \alpha_{j\mathcal{I}}, \eta_{j\mathcal{I}})$. The maximum likelihood variance of elements of θ_j:

$$\sigma_g^2 = \sum_{j=1}^{J} \pi_{jg}(z_{jg} - \sum_{j=1}^{J} \pi_{jg} z_{jg})^2 \tag{25}$$

where z_{jg} represents either the intercept term θ_{jg}, or the slope on parents rank η_{jg}, and $\pi_{jg} = \frac{n_{jg}}{N_g}$ are group-specific observation share , where n_{jg} is the number of children of group g residing during childhoods in city j for at least one year, and $N_g = \sum_{j=1}^{J} n_{jg}$. Note that we can write (25) also as:

$$
\begin{aligned}
\sigma_g^2 &= \sum_{j=1}^{J} \pi_{jg} z_{jg}^2 - \left(\sum_{j=1}^{J} \pi_{jg} z_{jg}\right)^2 \\
&= \sum_{j=1}^{J} (1-\pi_{jg}) \pi_{jg} z_{jg}^2 - 2 \sum_{j=1}^{J} \sum_{k=j+1}^{J} \pi_{jg} \pi_{kg} z_{jg} z_{kg} \\
&= S_{ML}.
\end{aligned}
$$

Let $z_g = (z_{1g}, ..., z_{Jg})'$. Then we can represent (25) as a quadratic form:

$$S_{ML} = z' \tilde{A} z$$

where

$$\tilde{A} = \begin{pmatrix} (1-\pi_{1I})\pi_{1I} & -\pi_{1I}\pi_{2I} & \cdots & -\pi_{1I}\pi_{2I} \\ -\pi_{2I}\pi_{1I} & (1-\pi_{2I})\pi_{2I} & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -\pi_{JI}\pi_{1I} & -\pi_{JI}\pi_{2I} & \cdots & (1-\pi_{JI})\pi_{JI} \end{pmatrix}.$$

To get an unbiased estimator for the variance, we multiply by $\frac{N_I}{N_I - 1}$:

$$S_U = \frac{N_I}{N_I - 1} S_{ML}$$

And therefore work with $A = \frac{N_I}{N_I - 1} \tilde{A}$, and get $S_U = z' A z$.

Respectively, the unbiased estimate of the covariance of z_j and z_j', elements of θ_j, can be written as:

$$\sigma_{zz'} = z' A z.$$

Lastly, note that using the representation above, we can represent any of the elements of Ω as a quadratic form as a function of θ. For example, the variance of $\alpha_{j\mathcal{N}}$ can be written as

$$\sigma_I^2 = \theta' B \theta$$

where

$$B = \begin{pmatrix} A & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and the covariance between $\alpha_{j\mathcal{N}}$ and $\alpha_{j\mathcal{I}}$ can be written as

$$\sigma_I^2 = \theta' B \theta$$

where

$$B = \begin{pmatrix} 0 & 0 & A & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since we do not observe θ but its noisy estimate $\hat{\theta}$ and its sampling variance Σ, an unbiased estimate for the variance component is therefore:

$$\widehat{\theta' B \theta} = \hat{\theta}' B \hat{\theta} - Tr(B\Sigma)$$

where $Tr(\cdot)$ is the trace operator.

F.2 Estimating the Sampling Variance of $\hat{\sigma}^2$

Our analytic standard errors of the variance component rely on the assumption that $\theta_j \sim N(\mu, \Omega)$, where $\mu = E[\theta_j]$. Therefore, for every two variance components $\theta' A \theta$, and $\theta' B \theta$, we estimate the sampling variance of the variance component Ω in the following way. First, we note that

$$Cov(\hat{\theta}' A \hat{\theta}, \hat{\theta}' B \hat{\theta}) = E[\hat{\theta}' A \hat{\theta} \hat{\theta}' B \hat{\theta}] - E[\hat{\theta}' A \hat{\theta}] E[\hat{\theta}' B \hat{\theta}],$$

Then, using the fact that

$$E[\hat{\theta}' A \hat{\theta} \hat{\theta}' B \hat{\theta}] = Tr(A\Sigma(B + B')\Sigma) + \theta'(A + A')\Sigma(B + B')\theta + E[\hat{\theta}' A \hat{\theta}] E[\hat{\theta}' B \hat{\theta}]$$

we get

$$Cov(\widehat{\hat{\theta}' A\hat{\theta}, \hat{\theta}' B\hat{\theta}}) = \hat{\theta}'(A + A')\Sigma(B + B')\hat{\theta} - 0.5Tr((A + A')\Sigma(B + B')\Sigma)$$

G The Effect of Early Migration

In Table 2, we find that among immigrant children, spending one more year in the average Israeli city increases immigrant children's income rank in adulthood. This raises the follow-up question: is this true not only for the average city but for all cities in Israel? Column (1) in Appendix Table G.1 provides a test for the null that the effects of all the cities in Israel are strictly positive by using Bai et al. (2022) testing procedure. We find that for both low and high-income families, we reject the null with over 99% confidence, which implies that although exposure to most of the cities generates higher income than exposure to the former Soviet Union, some cities would be relatively harmful to children in terms of long-run income rank. With a knowledge of the mean and the variance, and following our finding in Appendix Figure I.1 that location effects are approximately normal, we conclude in column (2) that 16-17% of the causal effect of cities is lower than the USSR's for low-income immigrants.[42]

Our findings that immigrating earlier is beneficial on average align with Alexander and Ward (2018)'s findings from the migration wave to the US at the end of the 19th century, which show using a within-siblings analysis that immigrating at a younger age had a positive effect on children's income in adulthood.[43] Our analysis extends these findings by analyzing the distribution of effects and providing evidence on the prevalence of gains. Nevertheless, we also show that even though the effects are mostly positive, there are cases in which immigrating one year earlier could result in adverse effects. This suggests that the determinants of successful early immigration depend not only on the sending country but also on the local environment within the destination country where the immigrant lives.

[42] In Appendix Figure I.1, we estimate the marginal distribution of each group location effect by employing the Efron (2016) log-spline estimator and find that a normal distribution well approximates it.

[43] In Appendix Figure A.4, we verify our results by studying the relationship between age at arrival and income rank at age 28 conditional on family fixed effects, therefore comparing siblings arriving in Israel at different ages.

Table G.1: Testing whether immigrating one year earlier to all the Israeli cities generates income higher than the USSR

	(1) P-val	(2) Share $\theta_j < 0$
θ_{25}	0.021	0.168
θ_{75}	0.003	0.165

Note: This table reports in column (1) the P-value from Bai et al. (2022) test that all the location effects of immigrants are always positive, and in column (2) the share of cities with negative location effect compared to the USSR.

H Neighborhood grouping

H.1 Neighborhood Definitions

We construct the geographic unit of neighborhoods based on the geographic unit "statistical area". Statistical areas are defined in the census as areas of 3,000-5,000 inhabitants, similar to a census tract in the US which is aimed to have around 4,000 inhabitants. This allows us to check for whether intra-city segregation drives our results of location effects heterogeneity, as described in Section 7. The city divisions, along with the 2008 statistical area codes that comprise them, are as follows:

- **Tel-Aviv:** North Tel Aviv (111-235), Center Tel Aviv (311-625), South Tel Aviv (811-947), Jaffa (711-747)
- **Haifa:** North-West Haifa and the shore (111-434), North-East Haifa (511-644), South Haifa (711-945)
- **Jerusalem:** East Jerusalem (611-741, 1411-1614, 2100-2999), Jerusalem Center (811-864, 1011-1044, 1211-1355), South Jerusalem (1111-1147, 1621-1644), North Jerusalem (111-543, 911-934)
- **Be'er Sheva:** Be'er Sheva Old (111-314), Be'er Sheva New (411-645)
- **Natanya:** Natanya East + Natanya North (111-355), Natanya South (411-534)
- **Petach Tiqva:** Petach Tiqva West (111-324), Petach Tiqva East (411-524)
- **Rishon Lezion:** Rishon Lezion East (111-427), Rishon Lezion West (511-625)
- **Ashdod:** Ashdod North (111-244), Ashdod South (311-434)

In less dense areas, we group localities according to their regional councils. Regional councils in Israel, or "Mo'atzot Ezoriyot" in Hebrew, serve as the administrative bodies for a group of smaller, geographically close communities, typically in rural settings. The overarching governance provided by regional councils enables us to include smaller villages and towns in our analysis as single units due to their shared administrative bodies, geographic proximity, and usually also education institutions. The regional councils that are included in our sample are the following: Upper Galilee, Mateh Asher, Emek HaYarden, Emek HaMaayanot, Gilboa, Jezreel Valley, Hof HaCarmel, Hefer Valley, Mateh Yehuda, Gezer, Be'er Tuvia, Eshkol, Merhavim, Misgav, Golan, Shomron, Mateh Binyamin, and Gush Etzion.

I The Joint Distribution of Location Effects

For evrey city j, let θ_g be the $J \times 1$ vector of location effects of group $g \in \{\mathcal{N}, \mathcal{I}\}$, with the corresponding $J \times 1$ vector $\hat{\theta}_g$ of estimated location effects. We assume that the estimated location effect follows a normal distribution: $\hat{\theta}_g \sim \mathcal{N}(\theta_g, \Sigma_g)$, which can be justified by central limit theorem with a growing number of families in each city. Our goal is to estimate the joint distribution of location effects to perform decision-making while allowing the mean location effects to vary linearly with a few city-level covariates z_j.

Abstracting from the p subscript for simplicity, our model for location effects is described by:

$$\begin{aligned} \theta_{jg} &= z_j' \beta_g + \nu_{jg} \qquad & \nu_j | z_j, \Sigma &\overset{iid}{\sim} G \\ \hat{\theta}_{jg} &= \theta_{jg} + u_{jg} \qquad & U_g | z_j, \Sigma &\sim \mathcal{N}(0, \tilde{\Sigma}) \end{aligned} \tag{26}$$

for $g \in \{\mathcal{N}, \mathcal{I}\}$, where $\nu_j = (\nu_{j\mathcal{N}}, \nu_{j\mathcal{I}})'$, $U_g = (u_{1g}, ..., u_{Jg})$, Σ is the $2J \times 2J$ sampling error covariance matrix with Σ_g on the diagonal and zeros in the off-diagonal, and z_j is a $p \times 1$ vector of city level covariates, including diversity, locality welfare expenditure per-capita, the log standard error of $\hat{\theta}$, and constant of 1 for the intercept with $z = (z_1, ..., z_p)$ the $J \times p$ design matrix. Under the model in Equation (26), the prior distribution of the demeaned location effects $\nu_{jg} = \theta_{jg} - z_j' \beta_g$ is independent of z_j and Σ. Therefore, to estimate the prior distribution of the location effects of immigrants and natives, it is sufficient to estimate $\beta = (\beta_{\mathcal{N}}, \beta_{\mathcal{I}})'$ and G, the distribution of ν_j.

Throughout this paper, we estimate the prior distribution of immigrant-native location effects by taking a two-step approach. First, we estimate β by city-size weighted least squares regression and form each group's $J \times 1$ residual $r_{jg} = \hat{\theta}_{jg} - z_j' \beta_g$ and covariance matrix $\tilde{\Sigma} = M \Sigma_g M'$, where $M = I - z(z'z)^{-1} z'$ is the corresponding projection matrix. Then, in the second step, we estimate the joint distribution of $r_j = (r_{j\mathcal{N}}, r_{j\mathcal{I}})'$.

I.1 Log-spline Estimator

We start by estimating the marginal distribution of each ν_{jg} nonparametrically using the empirical Bayes deconvolution estimator from Efron (2016). Under this approach, the prior distribution is assumed to belong to an exponential family, estimated flexibly by a fifth-order spline. The spline parameters are estimated via penalized maximum likelihood, where the log-likelihood is weighted by the total number of residents in each city. Following the approach taken in Kline et al. (2024), the penalization parameter is chosen to match the mean zero and method of moments variance component estimate of ν_j:

$$r_j' B r_j - Tr(G\tilde{\Sigma}) \tag{27}$$

described in Appendix Section F.

Appendix Figure I.1 plots the marginal deconvolved distribution of r_{jg} separately for every $g \in \{\mathcal{N}, \mathcal{I}\}$ and low and high-income families (solid blue line) together with the density of a normal distribution with the same mean and variance (dashed line). Results strongly suggest that each marginal distribution is well approximated by a normal distribution.

I.2 Normal Prior

Following the previous results, we assume that G follows a mean zero normal distribution with 2×2 covariance matrix Ω. With normally distributed signal and noise, the joint distribution of the estimated location effects is given by:

$$\hat{\theta}|\beta, \Omega, \Sigma, z \sim \mathcal{N}(\mu(x), V)$$

where $\mu(z) = (z\beta_{\mathcal{N}}, z\beta_{\mathcal{N}\mathcal{I}})'$, $V = \check{\Omega} + \tilde{\Sigma}$, $\check{\Omega} = \Omega \otimes I_J$, and I_J is $J \times J$ unit matrix. Lastly, the posterior distribution we exploit for decision-making is given by

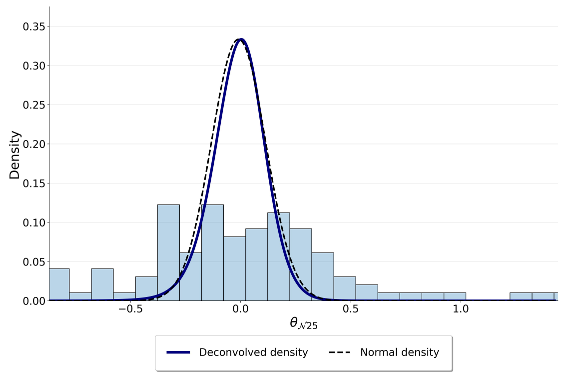
$$\theta|\hat{\theta}, \Sigma, \Omega, \mu_\theta(z), z \sim N(\theta^*(z), (\check{\Omega}_\nu^{-1} + \Sigma^{-1})^{-1})$$

where

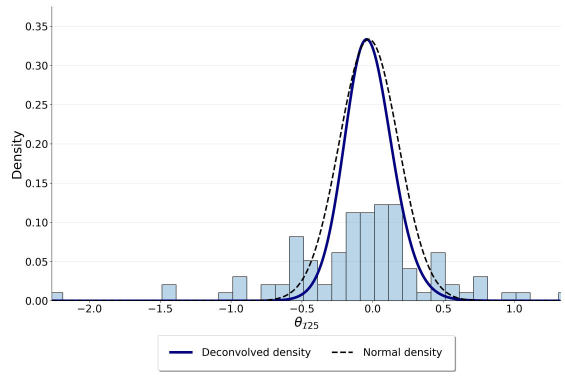
$$\theta^*(z) \equiv \mathbb{E}[\theta_{jg}|\hat{\theta}, z] = (\check{\Omega}_\nu^{-1} + \tilde{\Sigma}^{-1})^{-1}\left(\check{\Omega}_\nu^{-1}\check{\mu}_\theta(z) + \tilde{\Sigma}^{-1}\hat{\theta}\right),$$

and β and Ω are estimated via weighted least squares and method of moments as described above.

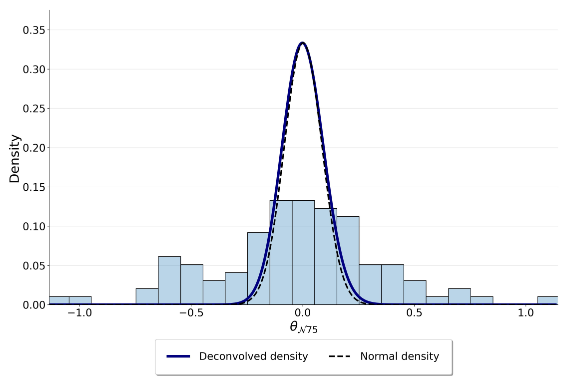
Figure I.1: Deconvolved density of childhood location effects of immigrants and natives



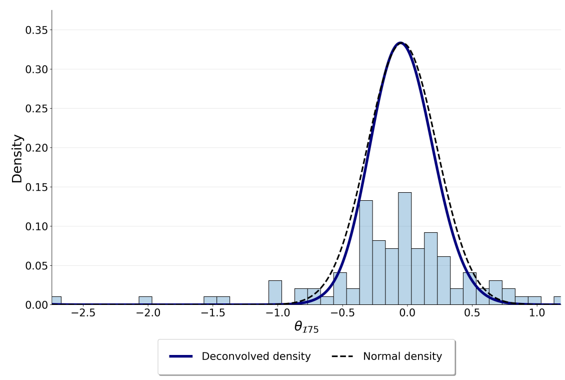
(a) θ_{25} Natives



(b) θ_{25} immigrants



(c) θ_{75} Natives



(d) θ_{75} immigrants

Note: These figures display the log-spline estimates of the distribution of the residuals of childhood location effects of immigrants and natives. The residuals are the difference between the estimated location effects $\hat{\theta}$ and the $z_j'\beta$, where β, presented in panel (i)-(iv) of Table A.4, is estimated in a city-size weighted least squers regression. Panel (a) presents the distribution for natives from the 25th percentile of the national income distribution. Panel (b) presents the distribution for immigrants from the 25th percentile of the national income distribution. Panel (c) presents the distribution for locals from the 75th percentile of the national income distribution. Panel (c) presents the distribution for immigrants from the 75th percentile of the national income distribution. The solid blue line shows the estimated deconvolved density following Efron (2016) penalized log-spline estimator with a natural cubic spline with five knots. The parameters of the deconvolved density were chosen to match the mean zero and variance from Table A.4. Histograms show the estimated location effects. Dashed lines show the density of normal distribution with the same mean and variance.

J Full List of Location Effects

Table J.1: Forcast of location effects for low-income families (p=25)

Loc. Code	Name	Posterior mean immigrants (1)	Posterior mean natives (2)	Share immigrants (3)
246	Netivot	0.431	-0.014	0.223
31	Ofaqim	0.417	0.007	0.299
7600	Akko	0.371	-0.044	0.277
4100	Qazrin	0.304	0.080	0.336
2660	Yavne	0.250	0.162	0.123
2630	Qiryat Gat	0.226	0.340	0.327
1034	Qiryat Mal'akhi	0.224	-0.066	0.220
1063	Ma'alot-tarshiha	0.220	0.224	0.479
73*	Mateh Binyamin	0.220	0.174	0.115
2560	Arad	0.219	0.165	0.419
2100	Tirat Karmel	0.207	-0.052	0.197
9100	Nahariyya	0.206	0.084	0.212
8500	Ramla	0.187	0.159	0.269
7700	Afula	0.181	0.100	0.302
6300	Giv'atayim	0.178	0.116	0.065
1139	Karmi'el	0.169	0.307	0.393
3640	Qarne Shomeron	0.169	0.203	0.155
7100	Ashqelon	0.166	0.167	0.351
5000	Tel-Aviv	0.163	0.077	0.127
1015	Mevasseret Ziyyon	0.151	0.185	0.111
8300	Rishon Leziyyon	0.147	0.310	0.187
2400	Or Yehuda	0.140	0.064	0.155
72*	Shomron	0.137	0.191	0.080
6200	Bat Yam	0.131	0.375	0.317
8700	Ra'annana	0.131	0.236	0.136
874	Migdal Haemeq	0.130	-0.079	0.289
1137	Qiryat Ye'arim	0.128	0.043	0.192
812	Shelomi	0.128	-0.422	0.202
99	Mizpe Ramon	0.118	0.085	0.245
3780	Betar Illit	0.113	0.189	0.087
9000	Be'er Sheva	0.106	0.180	0.306
76*	Gush Etzion	0.098	0.253	0.144
6600	Holon	0.097	0.304	0.177
1020	Or Aqiva	0.096	-0.011	0.437
240	Yoqne'am Illit	0.093	0.230	0.264
6400	Herzliyya	0.087	0.110	0.103
3650	Efrata	0.084	0.285	0.139
2800	Qiryat Shemona	0.079	0.024	0.183
7000	Lod	0.073	0.154	0.350
9600	Qiryat Yam	0.070	0.181	0.358
1066	Bene Ayish	0.066	0.285	0.581
6700	Tiberias	0.065	0.107	0.181
9400	Yehud	0.062	0.172	0.070
2500	Nesher	0.061	0.198	0.316
70	Ashdod	0.058	0.312	0.366
2530	Be'er Ya'aqov	0.058	0.148	0.147
9500	Qiryat Bialik	0.048	0.203	0.232
38*	Eshkol	0.048	0.112	0.075
4000	Haifa	0.046	0.304	0.272
1031	Sederot	0.040	0.110	0.392
8200	Qiryat Motzkin	0.035	0.344	0.220
8400	Rehovot	0.022	0.174	0.203
168	Kefar Yona	0.020	0.128	0.086

Continued on next page

Table J.1: Forcast of location effects for low-income families (p=25) (cont.)

Loc. Code	Name	Posterior mean immigrants (1)	Posterior mean natives (2)	Share immigrants (3)
3616	Ma'ale Adummim	0.014	0.303	0.149
1061	Nazerat Illit	0.007	0.122	0.521
1*	Upper Galilee	0.007	-0.141	0.134
26*	Mateh Yehuda	0.007	0.187	0.042
2034	Hazor Hagelilit	0.005	0.055	0.116
6100	Bene Beraq	-0.004	0.009	0.065
6900	Kefar Sava	-0.009	0.214	0.141
3570	Ari'el	-0.020	0.234	0.431
565	Azor	-0.040	0.161	0.105
831	Yeroham	-0.045	0.070	0.229
8600	Ramat Gan	-0.045	0.275	0.102
1224	Kokhav Ya'ir	-0.058	0.422	0.186
7900	Petah Tiqwa	-0.060	0.251	0.212
7400	Netanya	-0.061	0.236	0.274
7200	Nes Ziyyona	-0.065	0.129	0.083
3611	Qiryat Arba	-0.069	-0.213	0.186
2640	Rosh Haayin	-0.071	0.250	0.069
9700	Hod Hasharon	-0.084	0.304	0.071
681	Giv'at Shemu'el	-0.086	0.140	0.081
2620	Qiryat Ono	-0.095	0.163	0.071
6800	Qiryat Atta	-0.095	0.271	0.223
469	Qiryat Eqron	-0.096	0.212	0.124
8000	Zefat	-0.109	0.055	0.195
9300	Zikhron Ya'aqov	-0.110	0.202	0.077
2650	Ramat Hasharon	-0.122	0.219	0.038
2610	Bet Shemesh	-0.124	0.231	0.247
16*	Hefer Valley	-0.126	0.181	0.048
30*	Gezer	-0.126	0.219	0.030
6*	Emek HaYarden	-0.136	-0.061	0.099
71*	Golan	-0.136	0.201	0.059
2600	Elat	-0.136	0.336	0.180
9*	Jezreel Valley	-0.140	0.294	0.062
166	Gan Yavne	-0.144	0.121	0.108
7800	Pardes Hanna-karkur	-0.146	0.141	0.195
15*	Hof HaCarmel	-0.147	-0.031	0.119
56*	Misgav	-0.158	0.280	0.044
6500	Hadera	-0.163	0.376	0.303
4*	Mateh Asher	-0.165	0.174	0.000
2200	Dimona	-0.181	0.174	0.231
9200	Bet She'an	-0.182	-0.013	0.074
8*	Gilboa	-0.223	0.303	0.071
33*	Be'er Tuvia	-0.251	0.106	0.014
7*	Emek HaMaayanot	-0.251	0.176	0.052
2550	Gedera	-0.299	0.158	0.116
42*	Merhavim	-0.376	-0.268	0.000

Note: This table presents the posterior mean location effects on childrenwith parents' income from the 25th percentile. Columns (1) and (2) present the predicted location effects on immigrant and native-born children, accordingly. The table is sorted according to column (1). We list each location with its name and location code, where an asterisk marks regional council codes as detailed in Section H. Colum 3 presents the share of immigrants in the city in the year 2003.

Table J.2: Forcast of location effects for high-income families (p=75)

Loc. Code	Name	Posterior mean immigrants (1)	Posterior mean natives (2)	Share immigrants (3)
7600	Akko	0.542	-0.076	0.277
2400	Or Yehuda	0.466	-0.031	0.155
31	Ofaqim	0.413	0.001	0.299
246	Netivot	0.374	-0.084	0.223
72*	Shomron	0.362	0.007	0.080
4100	Qazrin	0.342	-0.148	0.336
6200	Bat Yam	0.338	0.267	0.317
3640	Qarne Shomeron	0.316	-0.045	0.155
1034	Qiryat Mal'akhi	0.287	-0.155	0.220
8500	Ramla	0.286	0.050	0.269
2660	Yavne	0.281	-0.015	0.123
7700	Afula	0.275	-0.014	0.302
9600	Qiryat Yam	0.271	0.049	0.358
2630	Qiryat Gat	0.267	0.168	0.327
812	Shelomi	0.226	-0.509	0.202
2100	Tirat Karmel	0.222	-0.223	0.197
5000	Tel-Aviv	0.212	-0.028	0.127
70	Ashdod	0.207	0.136	0.366
1066	Bene Ayish	0.203	-0.022	0.581
1015	Mevasseret Ziyyon	0.199	-0.016	0.111
73*	Mateh Binyamin	0.196	0.086	0.115
8300	Rishon Leziyyon	0.195	0.150	0.187
1020	Or Aqiva	0.194	-0.177	0.437
3650	Efrata	0.194	-0.037	0.139
6300	Giv'atayim	0.179	-0.006	0.065
1031	Sederot	0.178	-0.113	0.392
1137	Qiryat Ye'arim	0.175	-0.216	0.192
8700	Ra'annana	0.175	0.074	0.136
2600	Elat	0.173	0.129	0.180
2560	Arad	0.163	0.028	0.419
9400	Yehud	0.161	-0.050	0.070
1063	Ma'alot-tarshiha	0.155	-0.006	0.479
6600	Holon	0.145	0.179	0.177
2200	Dimona	0.138	-0.052	0.231
6700	Tiberias	0.132	-0.077	0.181
99	Mizpe Ramon	0.130	-0.097	0.245
9000	Be'er Sheva	0.094	0.022	0.306
2800	Qiryat Shemona	0.084	-0.205	0.183
76*	Gush Etzion	0.063	0.002	0.144
7000	Lod	0.059	0.030	0.350
7100	Ashqelon	0.058	0.034	0.351
240	Yoqne'am Illit	0.054	-0.004	0.264
6100	Bene Beraq	0.050	-0.021	0.065
6400	Herzliyya	0.050	0.009	0.103
7400	Netanya	0.029	0.076	0.274
9100	Nahariyya	0.024	-0.085	0.212
2640	Rosh Haayin	0.021	-0.010	0.069
38*	Eshkol	0.021	-0.270	0.075
2034	Hazor Hagelilit	0.007	-0.237	0.116
7900	Petah Tiqwa	0.006	0.087	0.212
2530	Be'er Ya'aqov	0.004	-0.109	0.147
1224	Kokhav Ya'ir	0.000	0.089	0.186
26*	Mateh Yehuda	-0.006	-0.085	0.042
2500	Nesher	-0.011	-0.012	0.316
30*	Gezer	-0.014	-0.038	0.030
8000	Zefat	-0.014	-0.155	0.195
8200	Qiryat Motzkin	-0.019	0.128	0.220

Continued on next page

Table J.2: Forcast of location effects for high-income families (p=75) (cont.)

Loc. Code	Name	Posterior mean immigrants (1)	Posterior mean natives (2)	Share immigrants (3)
166	Gan Yavne	-0.019	-0.062	0.108
9500	Qiryat Bialik	-0.023	0.044	0.232
874	Migdal Haemeq	-0.023	-0.218	0.289
1139	Karmi'el	-0.027	0.083	0.393
168	Kefar Yona	-0.028	-0.156	0.086
565	Azor	-0.029	-0.071	0.105
1*	Upper Galilee	-0.030	-0.341	0.134
3616	Ma'ale Adummim	-0.035	0.088	0.149
7200	Nes Ziyyona	-0.039	-0.099	0.083
2620	Qiryat Ono	-0.042	-0.092	0.071
2650	Ramat Hasharon	-0.047	-0.046	0.038
3780	Betar Illit	-0.051	-0.038	0.087
1061	Nazerat Illit	-0.061	-0.116	0.521
3611	Qiryat Arba	-0.064	-0.382	0.186
831	Yeroham	-0.067	-0.182	0.229
3570	Ari'el	-0.073	0.022	0.431
2610	Bet Shemesh	-0.098	0.166	0.247
71*	Golan	-0.100	-0.187	0.059
9300	Zikhron Ya'aqov	-0.100	-0.046	0.077
6800	Qiryat Atta	-0.103	0.034	0.223
6900	Kefar Sava	-0.111	0.048	0.141
6500	Hadera	-0.116	0.054	0.303
8600	Ramat Gan	-0.118	0.050	0.102
9700	Hod Hasharon	-0.126	0.035	0.071
9200	Bet She'an	-0.139	-0.355	0.074
7800	Pardes Hanna-karkur	-0.163	-0.046	0.195
4*	Mateh Asher	-0.164	-0.224	0.000
16*	Hefer Valley	-0.166	-0.052	0.048
8*	Gilboa	-0.168	-0.139	0.071
681	Giv'at Shemu'el	-0.184	-0.100	0.081
8400	Rehovot	-0.192	0.007	0.203
9*	Jezreel Valley	-0.196	-0.034	0.062
469	Qiryat Eqron	-0.200	-0.081	0.124
6*	Emek HaYarden	-0.220	-0.333	0.099
56*	Misgav	-0.225	-0.096	0.044
4000	Haifa	-0.253	0.086	0.272
15*	Hof HaCarmel	-0.258	-0.215	0.119
33*	Be'er Tuvia	-0.351	-0.196	0.014
7*	Emek HaMaayanot	-0.365	-0.224	0.052
2550	Gedera	-0.410	-0.161	0.116
42*	Merhavim	-0.479	-0.542	0.000

Note: This table presents the posterior mean location effects on childrenwith parents' income from the 75th percentile. Columns (1) and (2) present the predicted location effects on immigrant and native-born children, accordingly. The table is sorted according to column (1). We list each location with its name and location code, where an asterisk marks regional council codes as detailed in Section H. Colum 3 presents the share of immigrants in the city in the year 2003.

K Estimation of Minimax Decision Rules

We estimate several neighborhood recommendation policies in Section 9. In this appendix section, we provide further detail on how we estimate the policies of the minimax decision maker. We start by describing the benchmark models from Section 9.3 that put minimal restriction on location choices, and then turn to our extension from Section 9.6, which restricts to rational preferences that align partially with status quo sorting patterns.

K.1 Decision Rules Under Ambiguity

To account for the uncertainty of the decision maker regarding behavioral responses, we consider a minimax decision rule that seeks to be robust against the least favorable behavioral responses. First, the minimax decision rule in a model with uncertainty only regarding who shows up to each selected city is given by:

$$\delta_{jK}^{(\mathcal{N},\mathcal{I})} = \mathbb{1}\{\mathbb{E}[\max\{\vartheta_{j\mathcal{N}K}, \vartheta_{j\mathcal{I}K}\}|\mathcal{Y}] \leq \kappa_K\}, \tag{28}$$

which ranks locations based on their expected within-city posterior maximum regret, and where κ_K is the maximum value such that there are exactly K cities with $\mathbb{E}[\max\{\vartheta_{j\mathcal{I}K}, \vartheta_{j\mathcal{N}K}\}|\mathcal{Y}] \leq \kappa_K$. We refer this decision rule as *minimax over* $(\mathcal{N},\mathcal{I})$.

We consider a second minimax decision rule, which results from a model in which the decision maker faces uncertainty regarding who shows up and where they go. This decision rule has the following form:

$$\delta_K^{(\mathcal{N},\mathcal{I},\text{city})} = \arg\min_{\delta} \mathbb{E}[\max(\{\vartheta_{j\mathcal{N}K}, \vartheta_{j\mathcal{I}K}\}_{j\in S(\delta)})|\mathcal{Y}], \tag{29}$$

where $S(\delta) = \{j : \delta_j = 1\}$ is the set of recommended cities. Under this decision rule, the decisionmaker evaluates the posterior expectation of the maximum regret across all the selected locations and across immigrants and natives and chooses the list that attains the lowest worst-case regret. Therefore, hereafter, we refer to this policy as *minimax over* $(\mathcal{N},\mathcal{I},$*city*$)$.

Estimation - Estimation of the policies in Equations 28 and 29 requires computation of the posterior expectation of the maximum. While the posterior distribution of location effects is known and, therefore, can be used to compute each expectation by numerical integration, we reduce the dimensionality of our problem using the following identity. For simplicity, we present here the posterior expectation of the maximum across immigrants and natives, where the derivation below can easily generalized to the posterior expectation in Equation 29. let $W = \max\{\vartheta_{j\mathcal{N}}, \vartheta_{j\mathcal{I}}\}$. Then the CDF of W given $\mathcal{Y} = y$:

$$\begin{aligned} F_W(t|\mathcal{Y}=y) &= Pr(W \leq t|\mathcal{Y}=y) = Pr(\max\{\vartheta_{j\mathcal{N}}, \vartheta_{j\mathcal{I}}\} \leq t|\mathcal{Y}=y) \\ &= \int_t^{\infty}\int_t^{\infty} dG(\vartheta_{j\mathcal{N}}, \vartheta_{j\mathcal{I}}|\mathcal{Y}=y), \end{aligned}$$

Where $G(.|\mathcal{Y})$ is the posterior CDF of location effects, which, with a normal prior, follows a normal distribution. Therefore, the posterior expectation can be written as:

$$E[W|\mathcal{Y}=y] = -\int_{-\infty}^{0} F_W(t|\mathcal{Y}=y)dt + \int_0^{\infty}(1-F_W(t|\mathcal{Y}=y))dt.$$

where we compute the posterior maximum by plugging in $\hat{G}(.|\mathcal{Y})$, and commuting numerically $\hat{F_W}(t|\mathcal{Y}=y)$ and $E[W|\mathcal{Y}=y]$ over a one-dimensional grid.

K.2 Ambiguity with Restricted Location Choice Model

In Section 9.6, we consider a location choice model that aligns with the status-quo sorting shares $\pi_0 = (\pi'_{0j\mathcal{N}}, \pi'_{0j\mathcal{I}})'$ by bounding the maximum deviation from the status-quo spatial distribution. We measure the distance between the two distributions using the Total Variation distance function, which, together with the logical bound of $\pi_{jg}(\delta) \in [0,1]$, and the restriction of $\pi_{jg}(\delta) = 0$ if $\delta_j = 0$ we describe in Section 9.3, implies the following restrictions:

$$\pi_{jg}(\delta) \in [\max\{\tilde{\pi}^{\delta}_{jg0} - a, 0\}, \min\{\tilde{\pi}^{\delta}_{jg0} + a, 1\}], \qquad \text{with} \qquad \tilde{\pi}^{\delta}_{jg0} = \frac{\pi_{jg0}}{\sum_j(\pi_{j\mathcal{N}0} + \pi_{j\mathcal{I}0})\delta_j}, \tag{30}$$

for every $g \in \{\mathcal{N}, \mathcal{I}\}$, and for some $a \in (0,1)$. As a result, the implied minimax decision rule follows:

$$\delta^{\mathcal{N},\mathcal{I},city}_{K,R} = \arg\min_{\delta} \mathcal{R}^{\mathcal{N},\mathcal{I},city}(\delta),$$

Where

$$\mathcal{L}(\vartheta, \delta, \pi(\delta)) = \sum_{j=1}^{J} \delta_j(\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}K} + \pi_{j\mathcal{N}}(\delta)\vartheta_{j\mathcal{N}K})$$
$$\mathcal{R}_R^{\mathcal{N},\mathcal{I},city}(\delta) = \mathbb{E}[\mathcal{L}_R^{max}(\vartheta, \delta, \pi(\delta))|\mathcal{Y}]$$

Estimation - We estimate this decision rule by using a bootstrap implementation following Christensen et al. (2022). We start by drawing S draws of θ from the posterior distribution $\hat{G}(.|\mathcal{Y})$. Then, for each draw ϑ_s and for every δ such that $\sum_j \delta_j = K$, we solve the following linear programming problem:

$$\begin{aligned} \mathcal{L}_{R,s}^{max}(\vartheta_s, \delta) = &\max_{\pi(\delta)} \sum_{j=1}^{J} \delta_j(\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}Ks} + \pi_{j\mathcal{N}}(\delta)\vartheta_{j\mathcal{N}Ks}) \\ &\text{s.t} \quad \text{Equations 30}, \\ &\sum_j (\pi_{j\mathcal{N}}(\delta) + \pi_{j\mathcal{I}}(\delta)) = 1 \end{aligned}$$

and compute the bootstrap average maximum regret as

$$\mathcal{R}_R^{*\mathcal{N},\mathcal{I},city}(\delta) = \mathbb{E}^*[\mathcal{L}_R^{max}(\vartheta, \delta)|\mathcal{Y}] = \frac{1}{S}\sum_{s=1}^{S} \mathcal{L}_{R,s}^{max}(\vartheta_s, \delta). \tag{31}$$

Christensen et al. (2022) define the bootstrap efficient-robust decision rule as

$$\delta^{**\mathcal{N},\mathcal{I},city} = \arg\min_{\delta} \mathcal{R}_R^{*\mathcal{N},\mathcal{I},city}(\delta).$$

To avoid the estimating $\binom{J}{K}$ values of the expected maximum risks in Equation 31, we reduce the dimensionality of our simulation by taking a recursive approach. First, in step 1, we find the best $K = J - 1$ cities, which requires comparing only $J - 1$ values of bootstrap average maximum regret. Then, in step 2, to find the $J - 2$ optimal places, we search only among the optimal $J - 1$ locations, which implies computing only $J - 2$ values of 31. We continue this nested searching procedure until step i, for which $K = J - i$.

L A Restricted Choice Model

In this section, we mircofound the model of location choices presented in Section 9.6. Let $D_i \in \{1, ..., J\}$ be the random variable indicating the location choice of individual i to one of the J Israeli cities. We restrict attention to preferences that are consistent with the following choice model

$$D_i = \arg \max_{j \in \{1,...,J\}} U_{ij} - (a_{ij} + b_{ij}\delta_j), \qquad \text{for every } g \in \{\mathcal{N}, \mathcal{I}\}$$

where $U_i = (U_{i1}, ..., U_{iJ})$ is individual i's private evaluation, whose pdf is given by $f(u|g, \delta)$, conditional on immigration group $g \in \{\mathcal{I}, \mathcal{N}\}$ and policy δ. We do not require that U_i follows a specific distribution and allow U_{ij} and U_{ik} to be dependent for $j \neq k$.

The share of families of group $g \in \{\mathcal{I}, \mathcal{N}\}$ who choose to move to location j after observing recommendation δ can be written as:

$$\pi_{jg}(\delta) = \int \mathbb{1}\{U_j - (a_{gj} + b_{ij}\delta_j) \geq U_k - (a_{gk} + b_{ik}\delta_k) \text{ for all k}\} f(u|g, \delta) du,$$

where we set a_{gj} such that the location choice probabilities of individuals who face no recommendation (i.e., $\delta_j = 0$ for all j, represented at $\delta = 0$) equal to the status-quo sorting probabilities:

$$\tilde{\pi}^{\delta}_{jg0} \equiv \pi_{jg}(0) = \int \mathbb{1}\{U_j - a_{gj} \geq U_k - a_{gk} \text{ for all k}\} f(u|g, 0) du.$$

To see how this choice model is equivalent to the one in Section 9.6 let $\bar{b}_{gj}$ and $\underline{b}_{gj}$ be the lower- and upper-bounds of b_{ij} that satisfy the following conditions. If $\tilde{\pi}^{\delta}_{jg0} - a > 0$, $\underline{b}_{gj}$ satisfies

$$a = \tilde{\pi}^{\delta}_{jg0} - \int \mathbb{1}\{U_j - (a_{gj} + \underline{b}_{gj}\delta_j) \geq U_k - (a_{gk} + \underline{b}_{gl}\delta_k) \text{ for all k}\} f(u|g, \delta) du,$$

while if $\tilde{\pi}^{\delta}_{jg0} - a \geq 0$, $\underline{b}_{gj} \to -\infty$. Accordingly, if $\tilde{\pi}^{\delta}_{jg0} + a < 1$, $\bar{b}_{gj}$ is the value that satisfy

$$a = \int \mathbb{1}\{U_j - (a_{gj} + \bar{b}_{gj}\delta_j) \geq U_k - (a_{gk} + \bar{b}_{gl}\delta_k) \text{ for all k}\} f(u|g, \delta) du - \tilde{\pi}^{\delta}_{jg0},$$

and $\bar{b}_{gj} \to \infty$, otherwise.