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

Tslil Aloni Hadar Avivi*

October 11, 2025

Abstract

We study heterogeneity in the effects of childhood residential location on adult income for native-born Israelis and children of immigrants from the former Soviet Union. While place effects for high-income immigrants and natives are strongly correlated, for low-income families they are uncorrelated. Guided by this, we explore how this heterogeneity affects neighborhood recommendation policies that aim to recommend top Israeli locations without ethnicity-based targeting. Using empirical Bayes tools, we find that policies based on population-wide averages yield inferior outcomes for immigrants. Robust policies that guard against least favorable sorting patterns reveal 10 cities likely to benefit children in both groups.

*Hadar Avivi: University College London. Email: avivihadar@gmail.com. First version: November 2023. This version: October 2025. We thank 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, Aureo de Paula, Nick Flamang, Conrad Miller, Enrico Moretti, Jonah Rockoff, Ben Scuderi, Yotam Shem-Tov, Liyang Sun, and Damián Vergara. We thank seminar participants at UC Berkeley, UChicago, Yale, LSE, UCL, TAU, HUJI, Columbia GSB, SITE Conference, NBER SI Urban, and the OI Conference on Economic Mobility. We are grateful to Avigail Sageev for outstanding research assistance and to the Israeli Central Bureau of Statistics for invaluable administrative support. We are grateful to the Chief Economist at the Israeli Ministry of Finance 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 effect on outcomes in adulthood (see Chyn and Katz, 2021, for a review). This evidence is the basis for "moving to opportunity" policies that encourage low-income housing voucher recipients to move to high-opportunity neighborhoods (Katz et al., 2001; Bergman et al., 2019). These policies often provide a single unified recommendation about where people should move based on a ranking of population-wide neighborhood-level estimates. In the absence of prior knowledge about recipients' behavioral responses, the effectiveness of such unified policies relies on the assumption that location effects are characterized by a unique underlying ordering that limits heterogeneity. As evidence of heterogeneity grows (e.g., Chetty et al., 2018, 2020), there is increasing uncertainty about whether families who follow the policy recommendations will ultimately benefit.

In this paper, we study this question in two steps. First, we provide evidence that childhood location effects vary substantially for low-income children from different backgrounds. Using a comprehensive administrative dataset from Israel, we establish that, similar to Chetty and Hendren (2018a,b), one's place of birth contributes substantial variability to the adult earnings of both native-born and immigrant children. However, the correlation between these effects for low-income immigrant and native-born children is close to zero, suggesting that places that boost the income of one group do not necessarily benefit the other. Based on these findings, we then study the implications of this heterogeneity for the outcomes of the potential recipients of the neighborhood recommendation policies proposed in Bergman et al. (2019).

We begin by revisiting the benchmark estimates of childhood location effects from Chetty and Hendren (2018a,b) in Israel and separately estimate effects for immigrants from the former Soviet Union arriving in Israel between 1989 and 2000 and native-born children. Causal location effects are identified 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 random sorting, but rather assumes that among families with the same sequence of location choices, the child's age at arrival is unrelated to unobserved components that affect potential outcomes. We support this assumption with a series of robustness and specification tests.

Childhood location effects vary substantially for both native-born and immigrant children. For a child whose parents are at the 25th percentile of the income distribution, moving at birth to a one standard deviation better city boosts income at age 28 by 8-10% per year, compared to the mean. Childhood location effects also vary substantially within cities across immigration groups, with a pattern that differs by household income. 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 it is strong and positive 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. In contrast, our findings suggest that lack of social integration and assimilation—measured by the location effects on intermarriage rates—serves as a plausible explanation.

Large, diverse cities with a high immigrant share are more likely to benefit immigrants. This finding aligns with the literature, which emphasizes the effect of geographic concentration of immigrants and refugees on their outcomes (Edin et al., 2003; Beaman, 2012; Abramitzky et al., 2024). In contrast, municipality welfare expenditure per capita negatively correlates with low-income native-born location effect, while it is less predictive of low-income immigrant 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 the Russian immigrants in Israel.

Motivated by these findings, we next study the consequences of heterogeneity for the policy implemented in the Creating Moves 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), this restriction is motivated by legal and moral constraints common in many countries, where it is unacceptable to base public programs on ethnic identity or promote segregation.

Using a decision-theoretic framework, we start by evaluating the policy considered in Bergman et al. (2019), which ranks locations based on a pooled average estimate of city effect. By construction, a policy based on average effects puts lower weights on the gains for minority groups, producing inferior outcomes for such groups. These unequal outcomes arise from two sources: First, the decision-maker's inability to target the treatment by ethnic group ex-ante, which prevents the policy from leveraging the heterogeneity in location effects across groups, and 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, a hybrid Empirical-Bayes-*minimax* strategy, which provides a list of recommended locations that are optimal under the least favorable compliance scenario. We demonstrate that this policy can yield substantial benefits for minority groups and lead to more equitable outcomes. With this policy, we can pinpoint at least 10 cities that benefit both groups, in which the worst-case outcome for either group is 25% better than under the city-level average policy. Also, we can ensure that, on average, no more than 25% of the recommended cities would yield outcomes inferior to those resulting from the current status quo sorting patterns.

This paper contributes to several strands of the literature. First, we add a new perspective to the vibrant discussion on the challenges 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) and inference (Andrews et al., 2024; Mogstad et al., 2024), where the latter work emphasizes the ramifications of ranking locations based on noisy estimates rather than their true values. Although Chetty et al. (2018, 2020) acknowledge the potentially multifaceted nature of locations, the literature has not considered the complications it generates. As a result, analysis of both existing and proposed mobility policies behaves as if there is a single ladder of location effects. With treatment effect heterogeneity and a lack of knowledge about behavioral responses, there is no guarantee that families who comply with the policy will ultimately benefit. While Mogstad et al. (2024) note this concern regarding the risk of forming policy based on noisy estimates, similar logic also applies when the signal varies. This paper directly addresses this by modeling the uncertainty from both heterogeneity and unknown compliance, along with uncertainty driven by measurement error.

Methodologically, our work relates to a growing literature on Empirical Bayes (EB) 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, in which decisions result from constraints on various sorts of error rates (Gu and Koenker, 2020; Kline et al., 2022; Mogstad et al., 2024). 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. Similar to Christensen et al. (2022), our model departs from classic approaches to these problems (Manski, 2004) by solving a hybrid Bayes-minimax model that depends on point-identified parameters and 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. Following the EB paradigm, we estimate an EB-minimax model where we replace the subjective prior with the one we estimate from the data.

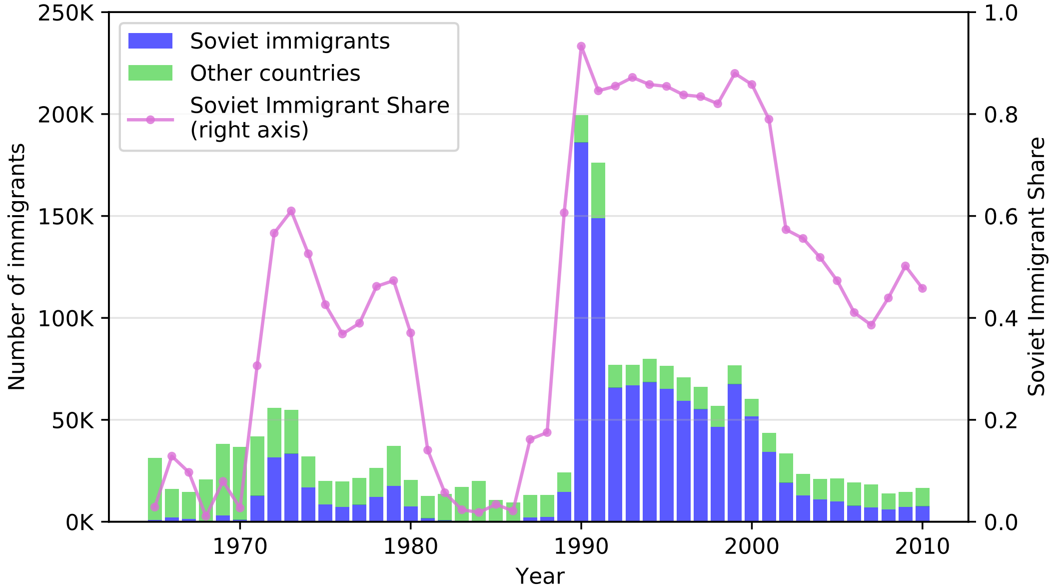
This paper also extends a growing literature in economics on algorithmic bias and fairness (Kleinberg et al., 2018; Cowgill and Tucker, 2019; Rambachan et al., 2020; Liang et al., 2021) and the equity-efficiency tradeoffs of affirmative action programs (Lundberg, 1991; Chan and Eyster, 2003; Ellison and Pathak, 2021). Papers in both strands often conclude that the optimal policy should exploit all available information, including group identity. In this paper, we explore the possibilities for a policy conditional on a suboptimal restricted algorithm, which, to our knowledge, hasn't been studied. 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 various other settings where policies face horizontal equity constraints.

2 Historical Context

In 1989, the Soviet Union relaxed its emigration restrictions, triggering 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

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



Note: This figure displays the number of migrants to Israel from 1965 to 2019 from the Soviet Union and other countries. Pink line (right axis) plots the fraction of Soviet immigrants.

of the migration wave, over 300 thousand immigrants, arrived in a relatively short time span, between 1989 and 1991, and accounted for 7% of the Israeli population prior to the immigration. This peak was followed by a steady influx of 60,000 per year throughout the decade, totaling over 1 million—one-fifth of Israel's 1989 population.

Soviet Jews received full Israeli citizenship upon arrival, granting them unrestricted access to social services, education, healthcare, and social security (Buchinsky et al., 2014). They faced no formal labor market restrictions and could settle anywhere in Israel. The government provided support, including a modest one-year grant ("absorption basket"), free Hebrew classes, and local integration centers.

This migration wave provides several favorable features for studying the effect of childhood location of residence on children's long-run economic outcomes. First, it was large and unrestricted, with entire families immigrating together, enabling us to estimate separately the effects for Soviet immigrants across multiple locations. Second, as citizens, immigrants faced no regulatory barriers compared to natives, ensuring that institutional factors can not explain any immigrant-native gaps.

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 $\sum_j e_j = 18$.[1] Throughout the paper, we assume that potential outcomes follow an additive structure:

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

where θ_{ij} represents the contribution to adult income of an extra year in city j to child i, and ξ_i is the error term, which includes all other age-, time-, or location-dependent shocks beyond the variation by exposure time and childhood city that affect children's long-run outcomes. It could include, for example, time-invariant and time-varying parental investments, moving costs, or age-specific shocks. This model rules out location effect heterogeneity by child's age, or complementarity or substitutability between time spent in different places. The observed outcome for child i is given by $Y_i = Y_i(E_i) = \sum_j \theta_{ij} E_{ij} + \xi_i$, where $E_i = (E_{i1}, ..., E_{iJ})'$ represents child i's realized years of exposure to each city from birth to age 18.

3.2 Identification Strategy and Research Design

The ideal experiment would randomly send children to different places at different ages. Absent such an experiment, we 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 native-born families that moved from city j to city l and are of the same income level, some children arrived at younger ages, and some arrived at older ages. Then, if among that narrow group, the moving decision is unrelated to the child's age at the move, we can infer the effect of growing up in city j compared with city l: $\theta_{jp} - \theta_{lp}$ by comparing

[1] Appendix Section D shows that, in Israel, moves after age 18 have little impact on children, consistent with the fact that most Israelis enlist in the army right after high school.

the outcomes of children who spent different time spans in each city.

Formally, consider all the families with the same family income rank $p(i)$ who moved once or twice between places when the child was younger than 18 years old, where $o(i)$ is child's origin location, which is always the USSR for immigrants, $d(i)$ is the destination location, and $d_2(i)$ is the second destination if the family moved twice and zero otherwise. We assume:

Assumption A1 *(Selection on observables)*

$$\xi_i \perp\!\!\!\perp E_i \mid (o(i), d(i), d_2(i), p(i)),$$

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. This includes no correlation with time-invariant factors, such as ability, or time-varying factors, such as parents' investments, that affect the child's income in adulthood and location choices. We test these assumptions in Appendix Section D. Importantly, Assumption A1 does not preclude systemic spatial sorting that correlates with the location effects themselves. 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.

3.3 Empirical Implementation

Building on this identification strategy, we estimate the childhood location effects of each city in Israel for children who moved between places in Israel when the children were young. Since our goal is to study heterogeneity by immigration group, we estimate location effects separately for immigrants and natives, with $g(i) \in \{\mathcal{N}, \mathcal{I}\}$ indicating whether child i is either native-born ($\mathcal{N}$) or immigrant ($\mathcal{I}$). To maximize sample size, we exploit variation in children's exposure time to different locations in Israel among families that experienced up to two moves when the child was young. For immigrants, it includes two groups. The first includes families that moved to Israel when the child was at age a_i, settled in city j, and stayed there until the child grew up. For these families, the exposure variable is $E_{ij} = 18 - a_i$ for the first city of residence j and zero otherwise. The second group consists of immigrants who moved twice, first immigrated to Israel when the child was at age a_i and settled in city $d(i)$, then moved to city $d_2(i)$ when the child was at age a_{2i}. For these families, exposure is given by:

$$E_{ij} = \mathbb{1}\{d(i) = j\}(a_{2i} - a_i) + \mathbb{1}\{d_2(i) = j\}(18 - a_{2i})$$

Similarly, for natives, our analysis includes families who moved once or twice between cities in Israel before the child turned 18, with a_i denoting the child's age at the first move from origin city $o(i)$ to destination city $d(i)$, and a_{2i} 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

$$E_{ij} = \mathbb{1}\{o(i) = j\}a_i + \mathbb{1}\{d(i) = j\}(a_{i2} - a_i) + \mathbb{1}\{d_2(i) = j\}(18 - a_{2i}).$$

Given these building blocks, we estimate the following OLS regression for children whose families moved between cities in Israel or immigrated to Israel before the child turned 18, separately for each immigration group:

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

where our main parameters of interest are the city-level slope coefficients on years of exposure, E_{ij}. We estimate heterogeneous location effects, allowing them to vary linearly by parental income rank, following earlier work indicating that a linear relationship between parental income rank and location effects provides a good empirical approximation (Chetty et al., 2014b).[2] 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 location j one-year return to parental income for a child who belongs to group g. Therefore, the total one-year location effect in city j for a child in group g with parental income p is θ_{jgp}.

In Equation (2), 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 level for immigrants.[3] By including the sequence of locations fixed effects, location effects are identified only from variation in the timing of moves rather than variation between families that moved between different places. Children's outcomes are measured at a fixed age and, therefore, across different calendar years. Therefore, we add the birth-cohort fixed effects to account for labor market fluctuations over time. Lastly, x_i includes birth year fixed effects interacted with parental income rank. We control for the sequences of locations and parental income in an additively separable way due to sample size constraints. Note

[2]Appendix Figure A.1 presents the relationship between children's income rank at ages 28-30 and parental income rank by immigration group and within a few selected cities, suggesting that this relationship is approximately linear in Israel as well.

[3]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.

that in this model, location effects are identified only in relative terms. Hence, we set immigrants' base-level location to be the former Soviet Union, and natives' base-level location to be Jerusalem.

Immigrants' origin location is grouped at the USSR level. Therefore, for this population, Assumption A1 requires the age of migration not to be correlated with the origin neighborhood within the USSR. This could be violated if the timing of when families left the Soviet Union varied across origin neighborhoods. In Appendix section D, we provide a list of robustness tests for the model's assumptions. In particular, we show that our specification tests are robust to the inclusion of family fixed effects, which, for immigrants, addresses the mentioned-above concern.

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.

3.4 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 finite-sample variance-covariance matrix of θ_j, denoted by Ω. For example, the variance of α_{jg} is

$$\sigma^2_{\alpha g} = \sum_{j=1}^{J} \frac{n_j}{N} (\alpha_{jg} - \sum_{l=1}^{J} \frac{n_l}{N} \alpha_{lg})^2, \tag{3}$$

where n_j is the number of families residing in city j, and $N = \sum_{j=1}^{J} n_j$.

We observe only noisy estimates of the location effects $\hat{\theta}_j$, rather than the location effects themselves, θ_j. Therefore, the sample variance, $\sum_{j=1}^{J} \frac{n_j}{N}(\hat{\alpha}_{jg} - \sum_{l=1}^{J} \frac{n_l}{N}\hat{\alpha}_{lg})^2$, 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 (3) is to subtract from the sample variance the mean squared of the standard errors (Chetty et al., 2014a; Chetty and Hendren, 2018b; Kline et al., 2022). As detailed in Appendix Section E, we use a variant of that estimator, which accounts for the correlation of the $\hat{\theta}_j$ across different j.

4 Data

We use administrative data collected by the Israeli Central Bureau of Statistics (CBS). The data covers the entire population of registered Israeli citizens born between 1950-1995 and their parents and comprises of four primary sources: tax records from the Tax Authority for the years 1995–2019 with employer-employee and self-employment tax information; education records from the Ministry of Education, including school identifiers and city; civil registry records providing demographics including gender, birth year, immigration date, and origin, family links (parents, siblings, spouse, children), and annual location of residence available for years 1999–2019 and 1995; and the 1995 censuse, which includes city of residence. The next section describes the sample construction and key variables. 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. Using the location of residence of both the child and the parents, we define the primary parent as the one who shares an address with the child for the majority of the years. If a location value is missing (for example, we have no information on the city of residence for years 1996-1998), we fill in the location of residence using the child's school location only if the school is in the same location as the child's location of residence in year $t-1$.[4] We enrich the location data using the city information available in the 1995 census. Specifically, we use the answers to two questions: "When did you move to your current city?" and "Where did you live 5 years ago?". Using these variables, we construct location information starting from 1995 and, for a subset, from 1990.[5] For the rest of this paper, our unit of location is a city or regional council,[6] which is the unit of local government.

For every parent in the sample, we construct the following variables: Parents' income, which is the total gross earnings at a household level, measured in 2016 Israeli shekels. In years when the family has no recorded earnings, the family's 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. We exclude families with less than 4 years of earnings, which accounts for 1.5% of parents. Finally,

[4] So differences between school locations and location of residence are not counted as moves.

[5] Response rate to these questions in 1995 is around 20%.

[6] A regional council is a group of nearby small towns or kibbutzim with a shared governing body.

we work with a parents' percentile rank variable, defined as the parental income rank in the national population that satisfies the restriction of having at least 4 years of earnings in 1995-2016. To account for the unbalanced structure of the child's age at which parents have earnings, we calculate each income rank within children's cohorts and, therefore, compare parents' earnings for children at the same ages. Children's main outcome is income rank at age 28, calculated within the child's cohort to account for differences in calendar year labor market conditions.

We study two primary populations. The first group consists of immigrants from the former Soviet Union who arrived in Israel between 1989 and 2000. We identify the children of immigrants based on their parents' birth country and year of immigration.[7][8] For each immigrant child, we calculate a_i, the age of the child when the family immigrated to Israel. We then code the first city or regional council of residence as their 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).[9] Similar to the immigrants, for every family we record all the cities in which the family lived during their children's childhood. In some cases, we refer to families who reside 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, we restrict attention to cities with at least 100 individuals in every group. These requirements narrow our analysis to 98 cities and regional councils out of 253.

4.2 Summary Statistics

Table 1 presents mean income (in Israeli shekels, \$1 $\sim$ 3.4 ILS) and income rank of parents and children of natives whose families moved between locations either once or twice and Soviet immigrants whose families 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. For a more detailed comparison of families that moved once and twice, see Appendix Section C.

[7] About 10% of immigrants in this period lacked country-of-origin data. We classified them as Soviet immigrants, since 90% of arrivals were from the former USSR (see Figure 1).

[8] For immigrants who arrived in Israel before 1995, who are the majority of immigrants, the data does not record the exact country of origin within the Soviet Union.

[9] Around 20% of Israelis are Arab citizens. However, geographic segregation—over 70% live in entirely Arab localities—limits our ability to compare the two groups.

Comparing columns 1-2 to 3-4, Table 1 shows that the 98 selected cities, which are Israel's largest cities and regional councils, represent most of the population, covering 88% of immigrants and 81% of native-born families that move between cities in Israel. Notably, immigrant parental income is 55% that of natives, echoing the results in Goldner et al. (2012) and Arellano-Bover and San (2023) of a large immigrant-native wage gap. However, by age 28, second-generation immigrants closed most of the gap, earning 95–96% of their native-born peers' income. This reflects high intergenerational mobility of immigrants, in line with findings in the US (Abramitzky et al., 2021).

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 and 2016 for immigrants and natives. All income variables are measured in Israeli shekels (1 US \$ $\approx$ 3.4 ILS). For immigrants, the sample includes all families who either arrived in Israel and stayed in the same city or arrived in Israel and then moved again between cities before the child turned 18. For natives, the sample includes all families that moved either once or twice between cities in Israel before the child turned 18. Columns 1-2 present the statistics for all families, and columns 3-4 present the statistics for families in our selected sample of 98 cities and regional councils.

Appendix Figure A.2 plots the geographic distribution of immigrants across Israel, both as their share of the total immigrant population (Panel (a)) and as their share within each locality (Panel (b)). As expected, the biggest cities like Haifa, Tel Aviv, Jerusalem, and Be'er Sheva absorbed the largest number of immigrants. However, Panel (b) illustrates that immigrants settled not only in large urban centers but also across the country, making up a significant portion of residents in many localities.

5 Estimates of Location Effects

5.1 Across-city Heterogeneity

Table 2 presents estimates of the distribution of location effects. As noted in Section 3.2, the cardinal value of location effects is not identified; therefore, for natives, they measure the effect of spending one more year in city j compared to one more year in Jerusalem, and for immigrants, they measure effects compared to one more year in the former Soviet Union. Panel (i) reports the mean and standard deviation of α_{jg} and η_{jg}. To summarize the full one-year effect of each city, panel (ii) reports the same statistics for $\theta_{jpg} = \alpha_{jg} + \eta_{jg} \times p$, for $p = 25$ and $p = 75$, which we refer to as the location effects of low- and high-income families, respectively. Standard errors for all variance estimates are based on the asymptotic variance, assuming that location effects are drawn from a normal distribution.[10]

Columns (1)–(3) report the statistics for all cities in Israel that meet the sample restrictions, separately by immigration group. In total, 153 cities and regional councils meet the criteria for natives, and 99 for immigrants. The average intercept α_{jg} of natives is 0.22, indicating that an additional year spent in the average city relative to Jerusalem boosts age-28 income for native-born children by 0.22 ranks. The corresponding estimate for immigrant children is 0.06 income ranks compared to staying one more year in the former Soviet Union. Note that because the estimates are in relative terms, η_{jg} can obtain negative values.[11]

Panel (ii) summarizes the distribution of location effects for low- and high-income families, separately by immigrant status. For both groups, mean location effects for families in the 25th percentile are positive, and for immigrants, they are statistically distinguishable from zero. For every year spent in the average city compared to the USSR (Jerusalem), low-income immigrant (native) age 28 earnings rank increases by 0.131 (0.143). A regression of children's income on their rank among children who spent their entire childhood in the same city shows that each percentile rank increase adds 1,530 shekels ($\approx$ \$450) for families at the 25th percentile and 1,689 shekels ($\approx$ \$490) for those at the 75th percentile. Using that relationship, the average city effect is equivalent to a 200 (219) Israeli shekels increase, which amounts to 60 (64) US dollars. 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 child's income for immigrants is always greater than zero, for low-income natives, the average city is better than Jerusalem, but for high-income, it's as good as Jerusalem.

[10] Using the log-spline estimator from Efron (2016), Figure A.5 estimates the marginal distribution of θ for immigrants and natives, which suggests that assuming normality is reasonable. Note that while the standard errors rely on normality assumptions, the estimates of the mean and variance are based on the method of moments and are therefore robust to functional form misspecification.

[11] We control for the interaction between parents' rank and children's birth cohorts. To test whether η_{jg} is always positive, we add the coefficients on parental income rank and birth cohorts to $\hat{\eta}_{jg}$ and apply the Bai et al. (2022) test. We reject the null that η_{jg} is always negative (p-values < 0.01) and cannot reject that it is always positive (p-values > 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)	153	0.223	0.257	98	0.214	0.245	177.7
		(0.113)	(0.044)		(0.115)	(0.042)	[0.0000]
Rank-parents (η)	153	-0.003	0.003	98	-0.003	0.003	179.3
		(0.001)	(0.001)		(0.001)	(0.000)	[0.0000]
Immigrants							
Cons. (α)	99	0.057	0.199	98	0.077	0.194	175.4
		(0.041)	(0.043)		(0.043)	(0.049)	[0.0000]
Rank-parents (η)	99	0.003	0.003	98	0.003	0.003	269.7
		(0.001)	(0.000)		(0.001)	(0.000)	[0.0000]
(ii) Total city effect							
Natives							
θ_{25}	153	0.143	0.236	98	0.137	0.200	166.9
		(0.103)	(0.036)		(0.105)	(0.037)	[0.0000]
θ_{75}	153	-0.018	0.247	98	-0.016	0.172	161.6
		(0.104)	(0.036)		(0.105)	(0.037)	[0.0001]
Immigrants							
θ_{25}	99	0.131	0.172	98	0.142	0.173	175.9
		(0.030)	(0.039)		(0.031)	(0.043)	[0.0000]
θ_{75}	99	0.280	0.211	98	0.273	0.218	247.9
		(0.033)	(0.030)		(0.035)	(0.032)	[0.0000]

Note: This table reports 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) report estimates for all cities; columns (4)-(6) restrict to cities with sufficient sample size for both groups. Columns (2) and (5) report the mean, and columns (3) and (6) the standard deviations, computed as the square root of the bias-corrected variance component. Panel (i) reports estimates for the distributions of the intercepts (α) and the slope coefficients on parental income rank (η). Panel (ii) displays the corresponding estimates for the distribution of location effects for children in the 25th and 75th percentiles of parental income distribution. Column (7) reports test statistics and p-values from chi-squared tests of the null hypothesis that all locations are identical. Standard errors for all variance estimates are based on the asymptotic variance, assuming that location effects are drawn from a normal distribution.

The standard deviations of θ_{jg25} and θ_{jg75}, presented in column (3), imply substantial across-city variation in location effects in Israel. For families at the 25th (75th) percentile, the standard deviation of θ_{jgp} is 0.236 (0.247) for natives and 0.172 (0.211) for immigrants—comparable to the magnitude of the standard deviation of the US county-level location effects estimates reported in Chetty and Hendren (2018b). Moving at birth to a city with a one standard deviation higher location effect for low-income natives (immigrants) increases children's income rank at age 28 by $0.23 \times 18 = 4.14$ ($0.17 \times 18 = 3.06$) ranks. This rank increase is equivalent to a 6,335 (4,681) ILS, or \$1,863 (\$1,377) increase, which is around 8-10% of the mean income of children with parents with below median income.[12] For comparison, the return to a matriculation certificate in Israel is 13% (Angrist and Lavy, 2009). Thus, moving at birth to a one standard deviation better city yields 61-77%

[12]Mean age-28 earnings for immigrants are 59,670 vs. 77,111 shekels (below vs. above median income); for natives, the corresponding figures are 57,884 and 73,448.

of the gains from earning this credential.[13]

Columns (5)-(6) report the same statistics for the 98 cities for which we estimate location effects for both immigrants and natives. For the remainder of the paper, we use this sample to study the heterogeneity in location effects in Israel. The estimates of the first two moments in columns (2) and (3) are not qualitatively different from those in columns (5) and (6), suggesting that this is not a special subset of cities. Finally, column (7) presents the χ^2 test statistic and corresponding p-value for the null hypothesis of no location effect heterogeneity across these cities. For all city-level parameters, we reject this null at conventional significance levels.

5.2 Immigrant-Native Differences in Childhood Location Effects

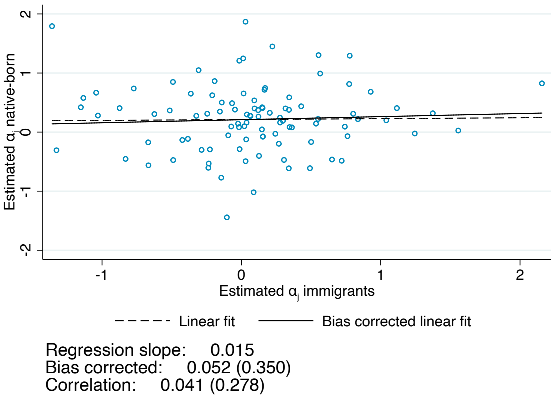
Location effects vary substantially between native and immigrant children. This is illustrated in Figure 2, which presents scatter plots and observation-weighted regression lines of the estimates of the effects for natives against the corresponding effects for immigrants, separately by income group. Figure 2a displays the relationship between immigrants' and natives' intercepts α_{jg}—i.e., between the location effects for 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 in the 25th and 75th percentiles of the national income distribution. Dashed lines are the naive attenuated regression lines, while solid lines are the biased corrected regression lines, with slopes estimated as the ratio between the covariance and the bias-corrected variance of immigrants' location effects. The corresponding estimates, together with the mean and standard deviation of within-city immigrant-native location effects gaps are shown in Table 3.[14]

The scatter plot and regression lines of the intercept (α_{jg}) in Figure 2a reveal 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 ($corr = 0.04$). In contrast, Figure 2b suggests much less heterogeneity in location effects as parental income increases. Places with high returns to parental income for immigrants tend to have high returns to parental income for natives. Combining these findings, Figure 2c shows no relationship between location effects for immigrant and native families at the 25th income percentile ($corr = -0.02$). However, for families in the

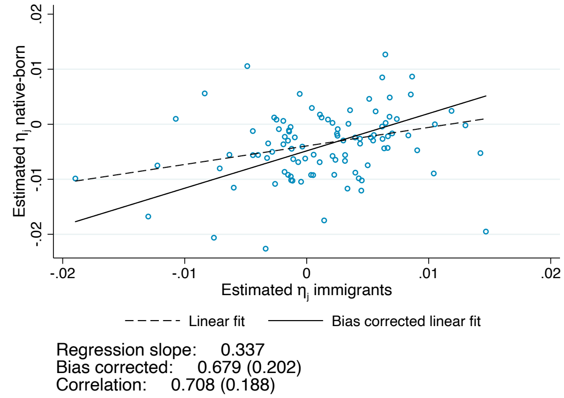
[13] High-school matriculation certificate is a key determinant of future labor market outcomes, as most post-secondary institutions require it for admission.

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

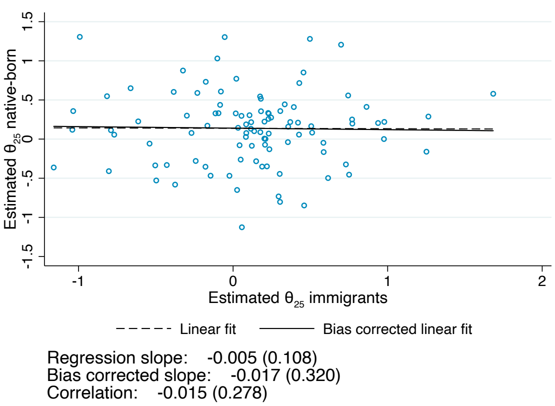
Figure 2: The relationship between location effects for immigrants and natives



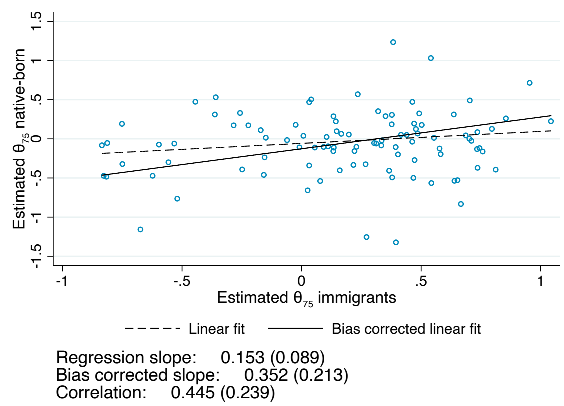
(a) α (intercept)



(b) η (slope)



(c) Total effect, θ_{25}



(d) Total effect, θ_{75}

Note: These figures display scatter plots and observation-weighted regression lines for immigrants' and natives' location effects. Panel (a) plots the estimated intercepts α_{jg}, panel (b) the estimated slopes η_{jg}, panel (c) the location effect for families at the 25th percentile of the income distribution, and panel (d) the location effect for families in the 75th percentile of the income distribution. Dashed lines show the naive regression line of $\hat{\theta}_{\mathcal{N}}$ on $\hat{\theta}_{\mathcal{I}}$ and the solid lines the bias-corrected regression line with slope $Cov(\theta_{\mathcal{N}}, \theta_{\mathcal{I}})/Var(\theta_{\mathcal{I}})$ using the estimates in Table 3.

75th percentile (Figure 2d), the location effects of immigrants and natives are strongly correlated. A correlation coefficient of 0.45 suggests that locations with one standard deviation higher effects for high-income immigrants have almost half of a standard deviation higher effects for natives.

The standard errors of 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 assuming normally distributed location effects. These intervals allow one-sided tests of whether each correlation coefficient equals 1. For low-income families—either at the bottom of the distribution or the 25th percentile—we can decisively reject correlations stronger than 0.5. In contrast, for η_{jg} (the return to parental income), the correlation is 0.70, and we cannot reject the null that it equals 1.

Table 3: Differences in location effects between immigrants and natives

	Covariance	Correlation	Implied OLS coefficient	Difference		
				Mean	Std.	χ^2 test $H_0 : \theta_{j\mathcal{N}} - \theta_{j\mathcal{I}} = c\ \forall j$
	(1)	(2)	(3)	(4)	(5)	(6)
α	0.002	0.041	0.052	-0.137	0.306	162.8
	(0.013)	(0.278)	(0.350)	(0.123)	(0.064)	[0.000]
		[-0.441, 0.538]	[-0.427, 0.590]			
η	0.000	0.708	0.679	0.006	0.002	130.5
	(0.000)	(0.188)	(0.202)	(0.001)	(0.001)	[0.016]
		[0.369, 1.000]	[0.383, 1.101]			
θ_{25}	-0.001	-0.015	-0.017	0.005	0.266	162.0
	(0.010)	(0.278)	(0.320)	(0.109)	(0.055)	[0.0001]
		[-0.514, 0.482]	[-0.500, 0.444]			
θ_{75}	0.017	0.445	0.352	0.289	0.209	146.8
	(0.009)	(0.239)	(0.213)	(0.111)	(0.053)	[0.0010]
		[0.010, 0.901]	[0.033, 0.799]			

Note: This table reports the relationship between immigrant and native location effects. Column (1) presents their covariance, column (2) presents the bias-corrected correlation, which is the covariance divided by the standard deviation of immigrants times the standard deviation of natives, 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 immigrants-natives gap, column (5) the standard deviation of the gap, and column (6) presents the chi-squared test statistics and p-values for the null of no location effect heterogeneity. First row reports estimates of the intercepts (α), the second reports the slope on parental income rank (η), and the last two rows report location effects for children in the 25th and 75th percentiles of parental income distribution. Standard errors of the variance and covariances are based on the asymptotic variance, assuming normally distributed location effects. Standard errors of the correlations and OLS slopes are calculated via delta method. Square brackets show the 90% parametric bootstrapped equal-tailed confidence intervals.

The last two columns of Table 3 report the mean and standard deviation of the immigrant-native location effect gap. Column (5) reveals substantial heterogeneity in the city-level location effects immigrant-native gap. The standard deviation of this gap is 0.27 and 0.21 for families at the 25th and 75th income percentiles—50%-30% higher than the standard deviation of the effects themselves. That is, moving at birth to a city with one standard deviation higher gap implies moving to a city that increases the adulthood income for one group by 7,436 ILS ($\approx \$2,181$) more than the other, which is more than 14% of the mean income at age 28 for children from a below-median-income family. Column (6) presents p-values for the null hypothesis of no within-city differences in location effects. In line with our findings, we can decisively reject the null of no within-city heterogeneity.

6 Predictors of Location Effects

Next, in Figure 3, we explore the characteristics of cities with high long-run effects on children's income by estimating the linear relationship between effects and city characteristics. Throughout this section, we demean the effects and the characteristics and divide them by the sample standard deviation. Most city-level characteristics come from early 2000s data, where detailed definitions and sources are in Appendix Section B. All regressions are weighted by city size.

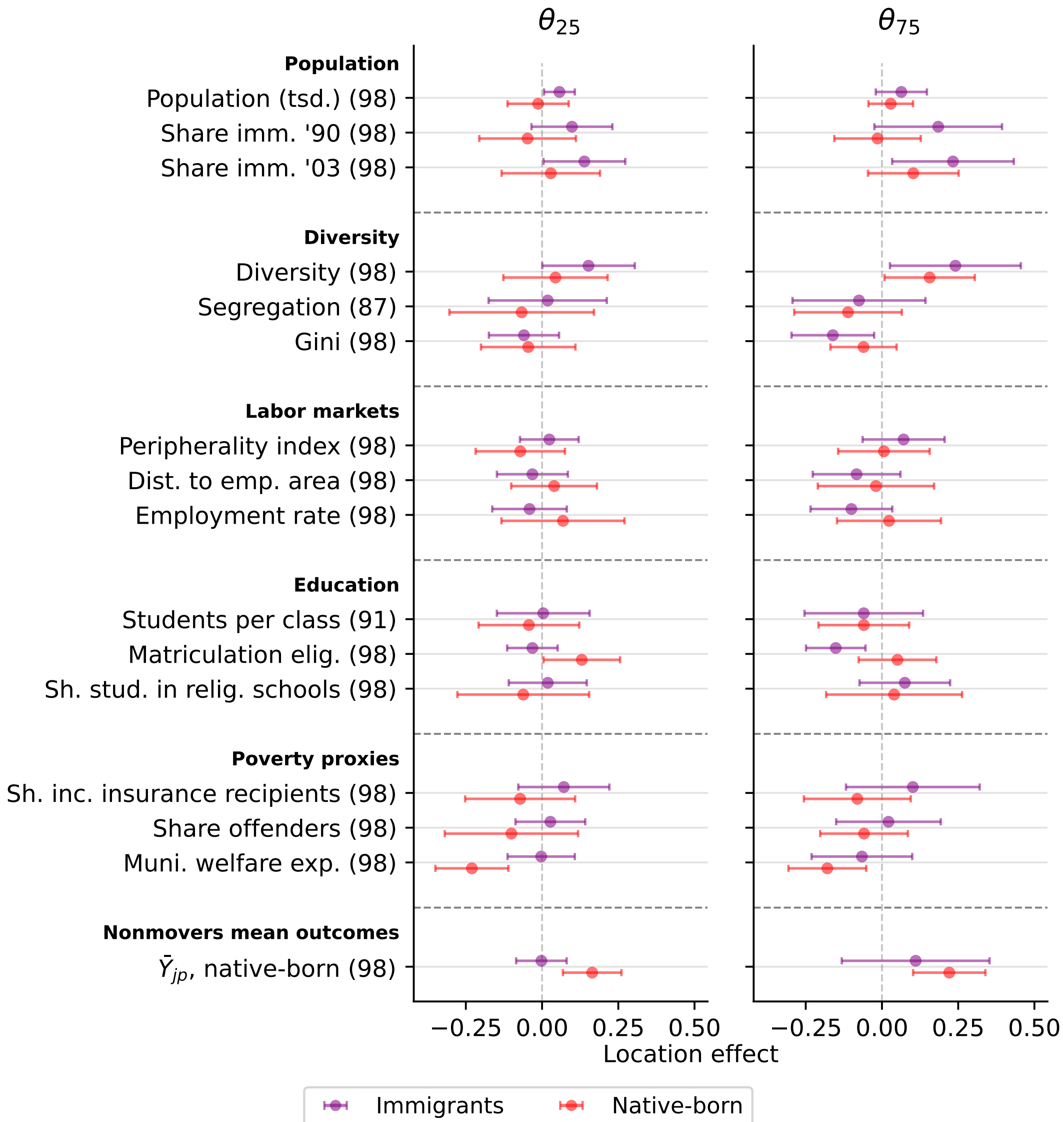
Population and diversity: The first rows in Figure 3 suggest that larger cities with a large immigrant share are associated with larger long-run effects on children of immigrants. The findings on the economic impact of the geographic concentration of ethnic groups on their outcomes are mixed. On the one hand, ghettos, mostly of the Black population in the US, have been found to have negative, lasting effects (Massey and Denton, 1993; Cutler and Glaeser, 1997; Chyn et al., 2022). In contrast, studies on refugees suggest a more nuanced relationship. Consistent with our findings, a handful of papers find that larger enclaves improve refugees' labor market outcomes through networks and social support (Edin et al., 2003; Beaman, 2012). Interestingly, in these papers, the city immigrant shares were at most 10%. In contrast, in a recent study of large Jewish enclaves in New York from the beginning of the 20th century, Abramitzky et al. (2024) find that Jewish immigrants who left the enclave saw earnings gains for themselves and their children. In that setting, the Jewish enclaves were huge, comprising over 60% of Jews. Inspired by this, we also estimate the relationship between location effects and diversity, measured with the entropy index which achieves its maximum value when city level immigrant share equals half and its lowest value when it is zero or one.[15] The fourth row in 3 suggests that city-level diversity positively predicts low-income immigrants' location effects.

A positive correlation between group shares and location effects could also reflect sorting, whereby 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 large 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.[16] Generally, our results call for more causal research to

[15]Diversity is: $-\left(\pi_{j\mathcal{I}}\ln(\pi_{j\mathcal{I}})+(1-\pi_{j\mathcal{I}})\ln(1-\pi_{j\mathcal{I}})\right)$, where $\pi_{j\mathcal{I}}$ is the city j immigrant share.

[16]Abramitzky et al. (2021) find somewhat different results for U.S. immigrants, showing they tend to live in areas with high mobility rates—likely reflecting high native-born long-run outcomes. However, their analysis is based on mean outcomes conditional on parental earnings, not causal location effects, so further evidence is needed for comparison.

Figure 3: Relationship between location effects and city characteristics



Note: This figure plots the relationship between city-level covariates and childhood location effects for children whose parents' income rank is in the 25th (left panel) and 75th percentile (right panel) of the income distribution. Each relationship is estimated with a bivariate weighted least squares regression, reweighting observations by population size in 2003, with location effects as outcomes. Covariates and location effects are standardized to have a mean of zero and a standard deviation of one in the sample. Bars indicate 95% confidence intervals based on robust standard errors. Appendix Section B describes the covariates' definitions. The number of cities in each regression is in parentheses. Cases with fewer than 98 localities are due to missing values or, in the case of segregation, because values cannot be calculated for cities that do not have sub-areas.

disentangle peer effects from sorting, with particular emphasis on the differences between immigrants and natives.

In contrast to evidence from the US, we find no relationship between low-income location effects and segregation, measured using Theil (1972) index and city Gini coefficient. Cities with higher income inequality are negatively associated with location effects for high-income immigrants.

Labor markets: Figure 3 finds that employment rates and proximity to employment centers and Tel Aviv, Israel's economic hub, are not predictive of location effects. One possible explanation for that is Israel's small size, with essentially one major employment center around the Tel-Aviv metropolitan area.[17]

Education: The next panel in Figure 3 studies the relationship between location effects and education inputs and outputs. For low-income families, cities with high rates of matriculation certificate attainment are associated with high location effects for natives but not immigrants. In Section 8, we investigate the role of high schools in further detail.

Poverty proxies: Figure 3 shows that municipality welfare expenditure per capita negatively predicts native-born children's location effects for all family incomes. The point estimates for the share of families receiving income insurance and the crime rate are also negative, although not precisely estimated. In our data, these are our best proxies for city poverty rates.[18] Interestingly, municipality welfare expenditure per capital is *not* predictive of immigrant location effects, further emphasizing the heterogeneity in our data. A negative relationship between location effects and poverty rates has also been found in the US and was one of the first measures the literature used for targeting housing policy (Katz et al., 2001). Their weak predictive power for immigrants suggests that using such a targeting policy would not be useful for immigrants in Israel.

City-level mean upward mobility rates: Previous research has emphasized that the observable mean child rank conditional on parental income rank for native born children is strongly predictive of location effects and suggests using these statistics for policy targeting. We demonstrate here that due to the high heterogeneity in location effects, it is not predictive of the benefits for all groups.[19][20] The last panel of Figure 3 shows that, in line with our evidence for heterogeneity, the native-born nonmovers' permanent residents' upward mobility rates, $\bar{Y}_{jp}$, are strongly predictive of natives' effects, with a point estimate similar to Chetty and Hendren (2018b). However, they have very little

[17] This is illustrated in Appendix Figure A.3, which plots the number of workers in each of Israel's employment centers, as measured in the 2008 census.

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

[19] We estimate upward mobility by regressing $Y_i = a_{j(i)} + b_{j(i)} p(i) + u_i$ for each city j, among native-born children whose parents never moved, where, $p(i)$ is parental income rank, and Y_i is the child's income rank at age 28. City j upward mobility rate is $\bar{Y}jp = \hat{a}j + \hat{b}_j p$.

[20] Note that because all the immigrants are included in our analysis, including those born in Israel, we can't compute the equivalent index for immigrants.

predictive power for low-income immigrants' location effects. These measures serve as the main instrument for guiding housing voucher policy in Bergman et al. (2019). Their weak predictive power for 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 10.

7 Possible Mechanisms

There are several plausible explanations for the lack of correlation between the location effects of immigrants and natives. For instance, immigrants and natives might attend schools of differing quality or reside in different neighborhoods within cities. Other possible explanations include the mismeasurement of immigrants' parental income or the possibility that the lack of income rank correlation reflects low rates of social interaction. In the following section, we directly test the first three explanations and find that they do not explain the zero correlation between the location effects of immigrants and natives. Although we are unable to fully disentangle the last explanation, we provide evidence suggesting that cities promoting social integration—measured by intermarriage rates—are more likely to exhibit high location effects for high-income immigrants but not for low-income immigrants.

High school fixed effects: We approximate the role of high schools in explaining variability in location effects by comparing variance components from Equation (2) with those from a regression that also controls for high school fixed effects.[21] To avoid dropping observations, we group high schools with fewer than five observations into one category.[22] This model is identified from cities with multiple schools and from schools that accept children from several local surrounding cities.

Appendix Table A.2 reports variance components for immigrants and natives with and without controlling for high-school fixed effects. The standard deviation of low-income native (immigrant) location effects declines from 0.20 (0.17) at baseline to 0.13 (0.10) when controlling for high-school fixed effects. Thus, high school effects explain $1 - \frac{0.13^2}{0.17^2} = 36\%$ ($1 - \frac{0.10^2}{0.17^2} = 41\%$) of the variation in location effects. However, the variance of the immigrant-native within-city gap is three times larger, and although the correlation coefficient becomes much noisier, the point estimate is more negative. This suggests that the zero correlation between immigrants and natives is not caused by schools. If anything, high schools in Israel act as equalizers. For high-income families, the drop

[21] We estimate fixed-effects rather than high school exposure effects due to computing limitations.

[22] As a result, there are 10 high schools in this grouped category.

in the correlation is even more striking, as, without the high school fixed effects, the correlation was strongly positive.

Neighborhood reweighting: The immigrant-natives heterogeneity might reflect differences in within-city sorting rather than heterogeneity in the effects themselves. To test that, one could estimate location effects at the neighborhood level and construct the city-level effects as the equally weighted average of neighborhood effects.[23] This approach was taken in Card et al. (2022) to estimate industry-level wage premia as the average firm effects. We follow a similar approach. We estimate Equation (2) 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.[24]

Appendix Table A.3 reports the results. Two key findings emerge. First, columns 1 and 3 show that the standard deviation of the reweighted estimates is 2 to 4 times larger than the unweighted estimates. Second, even after accounting for the differences in the spatial distributions of immigrants and natives, there is still substantial within-city heterogeneity where the correlation between the location effects of immigrants and natives remains zero for low-income families and strongly positive for high-income families. Similarly, in columns 2 and 4, we decisively reject the null of no heterogeneity. This suggests that although the differential spatial distributions matter for the magnitude of heterogeneity in location effects, they do not drive the disparities between immigrants and natives.

Parental income of immigrants not reflecting earnings potential: Immigrants often face earnings penalties due to frictions such as language barriers, cultural differences, and lack of networks and information. Cohen et al. (2001) document an occupational mismatch in which many Russian immigrants were college-educated but yet struggled to find jobs in their fields. 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 immigrant parents' income rank is lower than their skill or ability rank would suggest, we might mistakenly classify high-earning-potential parents as low earners. By doing so, when comparing the effects on immigrants and natives, we do not compare families with the same skills.

To accommodate this, Appendix Table A.4 reports the correlation matrix of high- and

[23] We can't estimate location effects at a neighborhood level due to our data agreement restrictions.

[24] Neighborhoods are based on statistical areas from the census for large cities (similar to census tracks in the US) and on the sub-villages for the regional councils.

low-income families, but instead of calculating parents' income rank in the national distribution, we calculate it within immigration groups and therefore rank parental income among comparable individuals. The standard deviation and correlations remain qualitatively similar to those in Table 2, suggesting that the zero correlation between low-income immigrants and natives is not due to misclassifying immigrants' income potential.

This test relies on the assumption that, despite occupational mismatch, within-group parental earnings remain correlated with skills, even if immigrants earn less than they would as natives. While intergenerational correlations are lower for immigrants than for natives (Aloni, 2017), they are still sizable and significantly different from zero, indicating that immigrant earnings are not entirely independent of their skills. Attenuated rank-rank correlations due to migration frictions are well-documented (Abramitzky et al., 2021). As long as this attenuation does not vary systematically across locations, it should not bias our cross-group correlation estimates.

The role of assimilation: A growing literature emphasizes the role of social interactions on children's long-term economic outcomes (e.g., Chetty et al., 2022). Therefore, a possible explanation for the lack of correlation in location effects among low-income immigrant and native families—and the high correlation among high-income families—might be a lack of social interaction and assimilation between the two groups among low-income families and higher assimilation rates among high-income families.

To test that, we approximate social assimilation with intermarriage rates of immigrants and natives, and study the relationship between cities that promote assimilation and their effect on income rank. As detailed in Appendix Section F, we proceed in two steps. First, we estimate the causal effect of each location in Israel on the intermarriage probability between immigrants and natives. In the second stage, we regress the income-rank location effects on the posterior mean estimates of intermarriage location effects. Interestingly, we find that while intermarriage effects are not predictive of immigrants' income-rank location effects for low-income families, they are predictive of high-income location effects. This result might suggest that a lack of assimilation and social interaction between low-income immigrants and natives could explain the zero correlation in low-income location effects on income rank.

8 Robustness and Research Design Validation

Research design validation: To support our identification strategy, Appendix Figures D.1 and D.2 provide a balancing exercise for the relationship between the age at move and age at arrival in Israel and parents' years of schooling, as measured in the 1995 census. For native-born children, we also estimate the relationship between parents' earnings growth when the child was young and the child's age at move. We find no statistically significant relationship between age at move and family characteristics conditional on location choices fixed effects and parental income rank.

Linear location effects: Our model assumes that location effects vary linearly with years of exposure, an assumption we test in Appendix Section D. Appendix Figure D.5 shows, similar to findings in other countries, a linear relationship between years of exposure and the mean outcomes of children who spent their entire childhood in the same location. D.7 validates this using childhood test scores, showing no effect for moves after the test age. Lastly, Appendix Figure D.8 shows that not only is the relationship between mean outcomes and exposure time linear, but also the relationship between location effects themselves and exposure time is well approximated by a linear function.

Change in natives' location effects: Derenoncourt (2022) finds that childhood location effects for incumbents could change due to a large migration wave. We test this by estimating Equation (2) separately for older cohorts born between 1980-1987 and younger cohorts born between 1988-1991.[25] Figure A.4 shows the correlation between the location effects of older and younger cohorts. Although cutting the sample in half increases sample uncertainty, point estimates suggest no change in location effects across the two groups.

Robustness checks and sensitivity: Appendix Section C estimates location effects using only one-time movers, yielding similar qualitative results despite the smaller sample size. Appendix Table A.5 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). Lastly, our current approach reweights cities based on the total number of families. Appendix Table A.6 shows that results are robust to reweighting cities 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 city-level group size.

9 The Distribution of Childhood Location Effects

Thus far, the paper describes the first two moments of the joint distribution of immigrant-natives' location effect, which imposes no restriction on the entire distribution of childhood

[25] Ideally, we would compare to cohorts unexposed to the Soviet migration wave, but no high-frequency micro-level locations data exist for those years.

location effects. Next, we extend our model and estimate this joint distribution. We use this extended model for two tasks. First, to form the posterior mean effect of each city, which is the best forecast of location effects that minimizes the mean squared error (James and Stein, 1961). Second, in Section 10, we exploit this distribution for a housing policy exercise, in which we generate predictions for other features of the joint distribution. In the following section, we briefly describe the model. For a more detailed discussion, see Appendix Section G.

9.1 Model and Estimation

We assume that the distribution of estimated location effects $\hat{\theta}_{jgp}$, and their population analog θ_{jgp} follows the following hierarchical structure:

$$\hat{\theta}_{jgp} = \theta_{jgp} + u_{jgp}, \quad U_{gp}|z, \Sigma_{gp} \sim \mathcal{N}(0, \Sigma_{gp}); \qquad \theta_{jgp} = z_j'\beta_{gp} + \nu_{jgp}, \quad \nu_{jp}|z_j, \Sigma_p \overset{iid}{\sim} \mathcal{N}(0, \Omega_p)$$

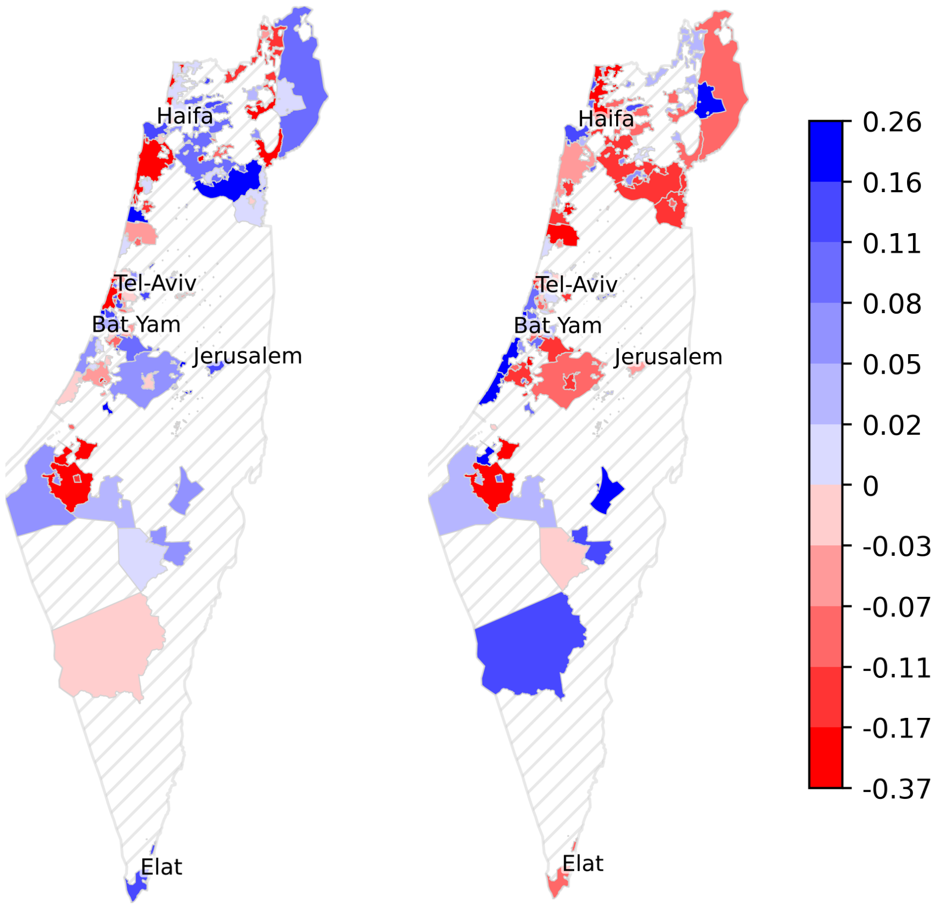
for $g \in \{\mathcal{N}, \mathcal{I}\}$ and parental income p, where $\nu_{jp} = (\nu_{j\mathcal{N}p}, \nu_{j\mathcal{I}p})'$, $U_{gp} = (u_{1gp}, ..., u_{Jgp})'$, Σ_p is the $2J \times 2J$ sampling error covariance matrix with Σ_{gp} on the diagonal and zeros in the off-diagonal, and z_j is a vector of city level covariates, discussed in Section 6, that were found to be the most predictive of location effect: city-level diversity index, population size, and locality welfare expenditure per capita (see discussion in Appendix Section G), with $z = (z_1, ..., z_J)'$. The normality assumption of $\hat{\theta}_{jgp}$ is motivated by the central limit theorem, while the normality assumption of $\theta_{jgp}|z_j, \Sigma$ is motivated by our finding in Appendix Section G that each marginal distribution is well approximated by a normal distribution.

We estimate the model in two steps. First, we run a weighted least squares regression of $\hat{\theta}_{gp}$ on z separately for every group $g \in \{\mathcal{N}, \mathcal{I}\}$. Then, similar to Section 3.4, we estimate the city-size-weighted unbiased variance component by method of moments, accounting for sampling error. Using the estimated hyperparameters (reported in column (2) in Appendix Tables G.1 and G.2) as prior, we estimate the posterior mean effect of each location for $p \in \{25, 75\}$, shrinking each $\hat{\theta}$ toward 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 squared error (James and Stein, 1961).

The location effects of low-income families: Figure 4 plots the demeaned posterior mean of location effects across cities and regional councils in Israel for immigrants and natives in the 25th percentile of the national income distribution.[26] The posterior means are highly variable both across cities and within cities across groups. The posterior mean

[26] Appendix Table H.1 lists the posterior means for all 98 cities and regional councils.

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



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

Note: These maps plot the posterior mean effects of year-long exposure to cities and regional councils in Israel on children's income rank at age 28 for children whose parents are in the 25th percentile of the national income distribution. Figure (a) displays the effects for natives and Figure (b) the effects for immigrants. The maps are constructed by grouping cities into 12 equally sized groups in which the darker blue the area, the larger its effect compared to the mean, and the darker red the area, the smaller the effect compared to the mean.

effect of spending one more year in the worst city ranged between 0.3 lower and 0.4 higher yearly income rank at age 28, a change of up to 620 ILS per year ($\approx$ \$180). Comparing immigrants and natives, it is apparent that there are significant differences between each group's city effect. Among low-income families, many of the northern Israeli communities are found to be places that benefit natives but not immigrants. In contrast, southern cities on the coastline of Israel, which have a high immigrant share, are among the best cities for immigrants, but are only as good as the average for natives.

10 Policy in the Face of Heterogeneity

Evidence on the importance of childhood residential location for 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. Katz et al. (2001) selected areas for public housing based on their poverty rates, 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, whereby the places that benefit low-income immigrants and natives are not necessarily the same places. Suppose we wanted to generate a list of recommended cities that provide the best opportunities for low-income children to inform housing policy in Israel, similar to Bergman et al. (2019). How does the treatment effect heterogeneity we document 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 used 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 Appendix Section K, we provide an extension with more realistic and elaborated assumptions.

10.1 Setup

Consider a decision-maker 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 target low-income families, we restrict attention to a policy that takes into account only the long-run effects on low-income children. As such, hereafter, to ease notation, we drop the parental income rank p subscript.

Our model incorporates several assumptions. First, due to ethical or legal considerations, the decision-maker is restricted to a decision rule that provides the same list of recommended cities to all groups, regardless of immigration status.[27] 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. For example, Bergman et al. (2019) prespecified a list of neighborhoods that promote upward mobility. Then, in an experiment on housing voucher recipients, they recommended the families in the treatment group to move to one of the neighborhoods on their list. Therefore, restricting the policy to be unified in this context implies precluding the

[27]See Appendix Section I for the legal motivation for this restriction, together with more examples of policies that impose horizontal equity constraints.

possibility of providing different recommendations to different groups. Another example could be a decision-maker who wants to choose K locations for new public housing units 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 decision-maker faces uncertainty regarding the true value of θ. However, she 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 decision-maker forms decisions by minimizing the expected, rather than the true, loss, where the expectation is taken over the posterior distribution of θ given $\mathcal{Y}$.

Lastly, we assume that the decision-maker acknowledges the loss that arises from banning the use of group characteristics and therefore evaluates the benefit of a decision rule relative to that expected from the first-best policy that allows for discrimination. Formally, let the *oracle* first-best policy for the best K cities for each group $g \in \{\mathcal{N}, \mathcal{I}\}$ be:

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

where $\theta^{(l)}_{jg}$ 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 $\theta^*(\delta^*_{gK}, K) \equiv \frac{1}{K}\sum_{j=1}^{J} \mathbb{E}[\theta_{jg}\delta^*_{jgK}]$ as the group g's expected long-run effect of selected cities under the first-best, where each expectation is taken over the distribution of the top K-th ordered statistics. Equipped with these definitions, the decision-maker values each city in comparison with the expected first-best value:

$$\vartheta_{jgK} = \theta^*(\delta^*_{gK}, K) - \theta_{jg}. \tag{4}$$

Equation (4) describes the *regret* of not using the first-best policy (Savage, 1951). With this normalization, the decision maker is concerned not only about the outcome she receives, but also about the outcome she would have received had she not been bound by ethical or legal constraints.

10.2 Benchmark: Selection Based on the Average Effect

We start with a model that rationalizes Bergman et al. (2019)'s selection criteria, in which the goal of the decision-maker is to choose the cities with the highest city-level average location effects. Formally, in Bergman et al. (2019), the decision-maker selects the list of cities, δ, that minimiazes the following loss function:

$$\mathcal{L}(\vartheta, \delta, \pi^0) = \sum_j \delta_j(\pi^0_{j\mathcal{I}}\vartheta_{j\mathcal{I}K} + (1 - \pi^0_{j\mathcal{I}})\vartheta_{j\mathcal{N}K}), \tag{5}$$

subject to $\sum_{j=1}^{J} \delta_j = K$, where $\pi^0_{j\mathcal{I}} = \frac{n_{\mathcal{I}j}}{n_{\mathcal{I}j}+n_{\mathcal{N}j}} \in [0,1]$ is the share of immigrants in city j in the data, i.e., in the status quo if no policy is enacted, and n_{gj} is the number of group $g \in \{\mathcal{N}, \mathcal{I}\}$ families in city j. This loss function implies that the decision-maker ranks places based on the pooled city-level mean effect, which describes how people sort within cities in the status quo:

$$\bar{\vartheta}_{jK} = \pi^0_{j\mathcal{I}} \vartheta_{j\mathcal{I}K} + (1 - \pi^0_{j\mathcal{I}}) \vartheta_{j\mathcal{N}K}, \tag{6}$$

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 regret, disproportionately 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 for immigrants.

The decision-maker does not observe the location effects directly. Instead, she knows their prior distribution and minimizes the expected loss—i.e., the Bayes risk—by choosing δ to minimize:

$$\begin{aligned} \mathcal{R}(\delta; \pi^0) &= \mathbb{E}[\mathcal{L}(\vartheta, \delta, \pi^0) | \mathcal{Y}] \\ &= \sum_j \delta_j (\pi^0_{j\mathcal{I}} \mathbb{E}[\vartheta_{j\mathcal{I}K} | \mathcal{Y}] + (1 - \pi^0_{j\mathcal{I}}) \mathbb{E}[\vartheta_{j\mathcal{N}K} | \mathcal{Y}]), \end{aligned}$$

subject to $\sum_{j=1}^{J} \delta_j = K$, where the expectation is taken over the posterior distribution of normlized location effects ϑ given the evidence $\mathcal{Y}$, and $\mathbb{E}[\vartheta_{jgK} | \mathcal{Y}]$ is the posterior mean regret of group $g \in \{\mathcal{N}, \mathcal{I}\}$ in city j. Therefore, the Bayesian decision-maker ranks locations by their posterior mean regret:

$$\mathbb{E}[\bar{\vartheta}_{jK} | \mathcal{Y}] = \pi^0_{j\mathcal{I}} \mathbb{E}[\vartheta_{j\mathcal{I}K} | \mathcal{Y}] + (1 - \pi^0_{j\mathcal{I}}) \mathbb{E}[\vartheta_{j\mathcal{N}K} | \mathcal{Y}],$$

and the optimal decision rule takes the following form:

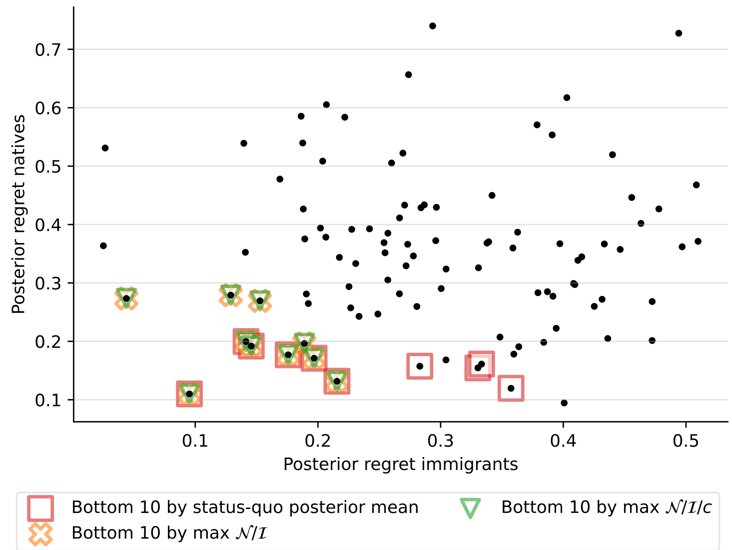
$$\delta_{jK} = \mathbb{1}\{\mathbb{E}[\bar{\vartheta}_{jK} | \mathcal{Y}] \leq \kappa_K\},$$

with κ_K being the value of the Kth lowest posterior mean $\mathbb{E}[\bar{\vartheta}_{jK} | \mathcal{Y}]$.[28] In what follows, we refer to this policy as targeting based on the *average status quo* sorting patterns.

Figure 5 displays a scatter plot of the posterior mean regret of immigrants and natives, where the red squares mark the cities selected under the average status-quo policy when $K = 10$. We can see that while the regret for natives is bounded below 0.2, the regret for immigrants could be almost twice as large.

[28]See Appendix Section K.1 for the proof.

Figure 5: Selected Cities Under the Average and Minimax Policies, $K = 10$



Note: This figure plots the scatter plots of immigrants' and natives' posterior mean regret for families at the 25th percentile of the income distribution. Black dots are the posterior mean regret for immigrants and natives. Red squares present the cities selected by the status-quo mean policy; Green rectangles present the cities selected by the minimax ($\mathcal{N}/\mathcal{I}/city$) policy; Orange Xs present the cities selected by the minimax ($\mathcal{N}/\mathcal{I}$) policy.

10.3 Accounting for Unknown Behavioral Responses

We now turn to explore alternative policies that strive to avoid harming any of the treated groups. In our setting, harm arises for two main reasons: the decision-maker's inability to offer personalized recommendations ex-ante, and the lack of information regarding which families will follow through and move to the recommended locations ex-post. Take-up uncertainty is a built-in restriction in the literature where all the proposed policies are primarily based on estimates of location effects, but not on estimates that incorporate information on recipients' compliance. This shortcoming was raised in Mogstad et al. (2024), who pointed 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. In this section, we propose a remedy by explicitly modeling the compliance uncertainty the decision-maker faces.

10.3.1 Who Shows Up?

We start with a simplified toy model where the decision-maker faces uncertainty only regarding the identity of the families that move to each recommended location. Such a scenario could arise, for example, if the decision-maker's task is to select a list of K cities for new public housing units, and build one unit in each chosen city. The loss function in this model mirrors Equation (5), but now the weights on each city represent 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, and responses to the policy, all of which are unknown.

Facing this uncertainty, the decision-maker can take several paths. Analogous to how the decision-maker 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 (6) could be that the decision-maker's prior reflects a belief that public housing recipients sort according to the status quo—that is, similar to existing sorting patterns within each city—regardless of the policy they face.

We opt for a different approach, acknowledging our ignorance regarding family sorting behaviors. We propose to follow a hybrid Empirical-Bayes-*minimax* policy, which is robust to the least favorable compliance scenario. Importantly, while we assume the decision maker confronts no prior knowledge on location choices, we maintain the assumption that the decision maker faces only statistical uncertainty about location effects, and can use $\mathcal{Y}$ to learn about their value. This leads to a policy, which we describe below, that minimizes the maximum regret over the set of all possible behavioral responses given ϑ (minimax), and then integrates over the distribution of $\vartheta|\mathcal{Y}$ (Bayes). This model builds on the theory of optimal statistical decision making (Manski, 2004; Kitagawa and Tetenov, 2018). However, in the spirit of Hurwicz (1951) and similar to Christensen et al. (2022), it departs from the traditional work by combining a minimax approach to handle ambiguity about compliance with expected risk minimization for the estimates of place effects.[29] The "Empirical" in our Empirical-Bayes-*minimax* policy, reflects the EB paradigm where, instead of using a subjective prior, we use the one estimated from the data.

Formally, given a vector of ϑ and a list of recommended cities δ, the worst-case regret that could have arisen under the least favorable compliance is:

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

where $\pi_{jg}(\delta) \in [0,1]$ decribes the probability that family of group g moves to city j given the list of selected cities δ, and $\pi(\delta) = (\pi_{1\mathcal{I}}(\delta), ..., \pi_{J\mathcal{I}}(\delta))'$. A minimax decision-maker who knows the true value of ϑ would directly seek to choose δ to minimize (7). However, with statistical uncertainty regarding the true value of ϑ, a Bayesian decisionmakers who

[29] Hurwicz (1951) proposed a generalized Bayes-minimax criterion that restricts the set of prior distributions in a minimax problem. In this paper, we restrict the prior of ϑ to a single distribution. A related Empirical Bayes minimax problem involving a restricted set of priors was introduced in Kline and Walters (2021).

knows the distribution of ϑ and the evidence $\mathcal{Y}$ chooses instead δ to minimize the following expected maximum loss:

$$\mathcal{R}^{(\mathcal{N},\mathcal{I})}(\delta) = \mathbb{E}\left[\mathcal{L}^{(\mathcal{N},\mathcal{I})}(\theta,\delta)\middle|\mathcal{Y}\right] = \sum_j \delta_j \mathbb{E}\left[\max_{\pi_{j\mathcal{I}}(\delta)} \{\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}K} + (1-\pi_{j\mathcal{I}}(\delta))\vartheta_{j\mathcal{N}K}\} \middle|\mathcal{Y}\right], \tag{8}$$

subject to $\sum_{j=1}^{J}\delta_j = K$, where the expectation is taken over the posterior distribution of ϑ given $\mathcal{Y}$. This Decision is motivated by optimizing an objective function involving both parameters that are set identified ($\pi(\delta)$) and parameters that are point identified (ϑ), sometimes referred to in the literature as robust Bayes decisions (Giacomini et al., 2021; Christensen et al., 2022).[30] This problem is equivalent to a zero-sum game with nature, where nature knows the true location effects and, for every list of recommended cities δ, it chooses the worst behavioral response $\pi(\delta)$. By minimizing the maximum *regret*, the decision-maker tries to achieve the optimal first-best solution without violating the horizontal equity constraint.

Minimizing the objective function in Equation (8) yields the following decision rule, in which the optimal policy is to rank locations based on their expected within-city posterior maximum regret:

$$\delta_{jK}^{(\mathcal{N},\mathcal{I})} = \mathbb{1}\{\mathbb{E}[\max\{\vartheta_{j\mathcal{I}K}, \vartheta_{j\mathcal{N}K}\}|\mathcal{Y}] \leq \kappa_K\}, \tag{9}$$

where κ_K is the maximum value such that there are exactly K cities with $\mathbb{E}[\max\{\vartheta_{j\mathcal{I}K}, \vartheta_{j\mathcal{N}K}\}|\mathcal{Y}] \leq \kappa_K$.[31] This decision rule can be thought of as the decision-maker's tolerance for bad outcomes due to sorting. Thus, Equation (9) gives an economically motivated decision rule when facing compliance uncertainty. When the posterior expected regret of the worst-case behavior falls below the cost κ_K, it is rational to recommend moving to that location. Throughout this paper, we refer to this policy as *minimax over ($\mathcal{N},\mathcal{I}$)*.[32]

Figure 5 displays the posterior mean regret for immigrants and natives of each city, where the green rectangles mark the cities selected under the minimax ($\mathcal{N},\mathcal{I}$) policy. Unlike the cities selected under the mean status-quo policy, the minimax decision rule identifies places that provide relative benefits (bounded regret) for both groups.

[30] The resulting decision rule is robust in the sense that it minnimizes the maximum regret over the set of possible behavioral resposense (π) given ϑ, and uses the the data $\mathcal{Y}$ to learn about the true ϑ.

[31] See Appendix Section K.1 for the proof.

[32] For brevity, we hereafter refer to this hybrid Empirical-Bayes-minimax policy as the minimax ($\mathcal{N},\mathcal{I}$) policy, emphasizing that the regret maximization applies to the identity of the recipients ($\mathcal{N},\mathcal{I}$) but not to the location effects.

Connection to welfare economics: The spectrum of objectives between that implied by the observation-weighted loss function in Equation (5) and the minimax loss function in (7) maps to the familiar social welfare criteria. At one extreme, Equation (5) 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 decision-maker's social welfare weights. At the other extreme is the Bayes-minimax decision rule in Equation (7), which is equivalent to a Rawlsian decision-maker 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 decision-maker's attitudes towards equity.

10.3.2 Who Shows Up and Where Do They Go?

Next, we consider a decision-maker who faces uncertainty not only regarding the identity of each housing recipient but also regarding their location choices. This model is directly inspired by Bergman et al. (2019)'s experiment and Mogstad et al. (2024)'s critique. In this framework, the decision-maker 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 (and with a slight abuse of notation), 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 decision-maker seeks to minimize the following loss function:

$$\mathcal{L}(\vartheta, \delta, \pi(\delta)) = \sum_{j=1}^{J} \Big[\delta_j(\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}K} + \pi_{j\mathcal{N}}(\delta)\vartheta_{j\mathcal{N}K}) + (1-\delta_j)(\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}K} + \pi_{j\mathcal{N}}(\delta)\vartheta_{j\mathcal{N}K}) \Big] \tag{10}$$

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 on the list. We justify this approach following the findings of the CMTO experiment. First, the CMTO experiment developed a technology that induces substantial compliance, which increased the share of families moving to recommended places by more than 38%. Second, Bergman et al. (2019) also find that the sorting pattern of the CMTO experiment control group aligns with the sorting pattern in the status quo, absent the experiment. Therefore, the regret associated with the second part of Equation 10 is likely constant, consisting of the share of noncompliers and the regret from the status quo sorting.[33] Therefore, we restrict attention to a decision-maker who minimizes the following loss function:

$$\mathcal{L}(\vartheta, \delta, \pi(\delta)) = \sum_{j=1}^{J} \delta_j (\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}K} + \pi_{j\mathcal{N}}(\delta)\vartheta_{j\mathcal{N}K}), \tag{11}$$

subject to $\sum_{j=1}^{J} \delta_j = K$. Since the decision-maker has no prior knowledge on $\pi(\delta) = (\pi_{1\mathcal{N}}(\delta), \pi_{1\mathcal{I}}(\delta), ..., \pi_{J\mathcal{N}}(\delta), \pi_{J\mathcal{I}}(\delta))'$, she seeks a robust policy, which is optimal even under the worst compliance pattenrs. For a given list of cities δ and normlized location effects ϑ, the loss associated with such sorting behaviour is:

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

where the policy compliers belong to the immigration group with the highest regret and sort into the worst recommended city. With uncertainty also regarding the true location effects, the robust Bayesian policy aims to select the δ which minimizes this expected loss:

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

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 cities on that list and across all groups:

$$\delta_K^{(\mathcal{N},\mathcal{I},\text{city})} = \arg\min_{\delta} \mathbb{E}[\max(\{\vartheta_{j\mathcal{N}K}, \vartheta_{j\mathcal{I}K}\}_{j\in S(\delta)})|\mathcal{Y}], \tag{12}$$

where $S(\delta) = \{j : \delta_j = 1\}$ is the set of recommended cities.[34] Under this decision rule, the decision-maker evaluates the posterior expectation of the maximum regret across all selected locations and across immigrants and natives and chooses the list with the lowest regret. Therefore, hereafter we refer to this policy as *minimax ($\mathcal{N}$,$\mathcal{I}$,city)*.

Figure 5 displays the posterior mean regret for immigrants and natives in each city, where the orange Xs are the cities selected under the minimax $(\mathcal{N}, \mathcal{I}, city)$ policy. In our setting, the cities selected under the minimax $(\mathcal{N}, \mathcal{I}, city)$ decision rule happen to be the same cities as those selected under the minimax $(\mathcal{N}, \mathcal{I})$ policy. Similarly, it identifies the places that provide relative benefits (bounded regret) for both groups.

[33]Formally, if we denote $\delta = 0$ be the no-policy where no recommendation is made and represent the regret from the no-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 out of the entire population. Then, the loss function 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, non-compliers will not affect the optimal policy.

[34]See Appendix Section K.2 for the proof.

10.3.3 Estimation

Regret and prior distribution: ϑ_{jgK} depends both on the distribution of θ_{jg}, which we estimate directly, and on $\theta^*(\delta^*_{gK}, K)$, which is a function of the distribution of the K-th ordered statistic. As detailed in Appendix K, for every K we approximate $\theta^*(\delta^*_{gK}, K)$ via Monte Carlo simulation with 10,000 draws from the mixing distribution of θ in column (2) of Appendix Table G.1. The estimates of the regret values $\hat{\vartheta}_{jgK}$ are then the difference between $\hat{\theta}_{jg}$ and $\theta^*(\delta^*_g, K)$, where $\theta^*(\delta^*_g, K)$ is treated as contant measured with no noise. Then, given $\hat{\vartheta}_{jgK}$, and their variance covariance matrix Σ, we estimate the prior distribution of ϑ_{jgK} following the same steps depicted in Section 9.

Posterior max policies: As detailed in Appendix K, the EB posterior expectations of the minimax policies are computed via simulation, drawing from the posterior distribution of $\vartheta|\hat{\vartheta}, \Sigma, z$. While most applied EB papers focus on posterior means of individual latent parameters, this paper—similar to Kline et al. (2024)—estimates EB posterior means of non-differentiable objects. To assess the sensitivity of our nonsmooth estimates to finite sample error, we provide in Appendix Section G.3 a robustness exercise using a full hierarchical Bayes model.

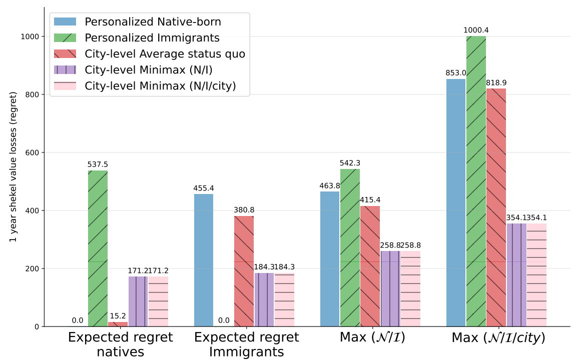
10.4 Evaluation of Each Policy

We evaluate each of the policies mentioned above via a Monte Carlo simulation that conditions on the noise structure Σ and the city covariates z of the 98 Israeli cities. At every simulation, we draw a new vector of $\vartheta|z$ and $\hat{\vartheta}|z$, and for a grid of values of K, we estimate the mean status-quo average policy and the minimax policies. We report the expected (true) mean regret for each group and the expected maximum regret for both groups and cities. See Appendix Section K for more details.

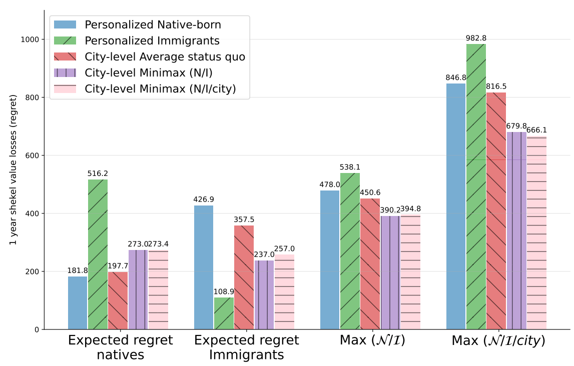
10.4.1 Counterfactuals

Figures 6a and 6b illustrate the costs and benefits of the policies described above for $K = 10$. Values are the regret shekel value associated with an additional year of exposure. That is, they reflect the expected money lost from not residing in the group-specific first-best cities for one year. The first two groups of bars report the expected average regret of immigrants and natives from selected cities: $\mathbb{E}[\frac{1}{K}\sum_j \vartheta_{jgK}\delta_j]$; the third, $\max(\mathcal{N}/\mathcal{I})$, reports the expected within-city maximum regret: $\mathbb{E}[\frac{1}{K}\sum_{j=1}^{J}\max\{\vartheta_{j\mathcal{N}K}, \vartheta_{j\mathcal{I}K}\}\delta_j]$ across immigrants and natives in the average selected city; and the forth, $\max(\mathcal{N}/\mathcal{I}/city)$, reports the expected maximum regret across all selected cities, immigrants, and natives: $\mathbb{E}[\max\{\vartheta_{j\mathcal{N}K}, \vartheta_{j\mathcal{I}K}\}_{j\in S(\delta)}]$. In subfigure 6a, we evaluate these policies for an oracle decision-maker who knows ϑ. Therefore, any regret reported there is only a consequence of the heterogeneity and the horizontal equity restriction. In subfigure 6b, we consider a decision-maker who doesn't know ϑ but relies on our estimates $\mathcal{Y}$, as described in the previous section. Therefore, the regret associated with each policy in this plot is a result of both heterogeneity and noise uncertainty.

Figure 6: Targeting trade-offs from choosing the top 10 cities



(a) Known θ (oracle)



(b) Unknown θ

	% better than the status-quo	
	Known θ (oracle) (1)	Unknown θ (2)
i) Personalized policy		
Native-born	0.548	0.534
Immigrants	0.398	0.414
ii) City-level policy		
Av. status-quo	0.643	0.584
Minimax ($\mathcal{N}/\mathcal{I}$)	0.999	0.745
Minimax ($\mathcal{N}/\mathcal{I}$/city)	0.999	0.733

(c) Share better than the status-quo

Note: These plots present the expected regret from policies selecting the top 10 places in Israel based on location effects for children with parents at the 25th income percentile. Regret, reported in shekels (1 US \$ $\approx$ 3.4 ILS), is the difference between the one-year location effects and the expected benefit from the top 10 cities for each group. The policies in subfigure (a) are based on the true location effects, and in subfigure (b), they are based on the empirical Bayes posterior means. The first two groups of bars report each group's expected regret. $Max(\mathcal{N}, \mathcal{I})$ bars report the expected within-city immigrant-natives maximum regret of selected cities, and $Max(\mathcal{N}, \mathcal{I}, city)$ bars report the expected maximum regret across all selected cities and immigrants and natives. Blue and green bars report policies that rank locations by group-specific regret; red bars report policies that rank places based on city-average regret (Eq. 6); purple and pink bars report minimax policies based on within-city and cross-city max regret (Eqs. 9 and 12, respectively). Subfigure (c) shows the share of selected cities with regret below the status quo.

As a result of the lack of correlation between the location effects of immigrants and natives, policy recommendations based on the effects on one group generate substantial regret for the other. This can be seen in the blue and green bars, which plot the outcomes from the personalized first-best policy in which the decision-maker recommends the top 10 locations based on the regret of only one of the groups, either immigrants or natives.

By construction, under full information on location effects (subfigure 6a), 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 537.5 ILS ($\approx$ \$158) lower income in adulthood compared with the first-best. Similarly, sending native-born children to the top 10 immigrant places provides them with 455.4 fewer shekels ($\approx$ \$133) in adulthood per year, compared to the first-best.

When ϑ is unknown (subfigure 6b), the policies are based on empirical Bayes shrinkage of functions of location effects. With the personalized policies, places are ranked according to the posterior mean effects of each group. The blue and green bars show that for immigrants (natives), the average recommended location under that personalized policy generates 108 (181) fewer shekels in adulthood than the group's first-best. This loss is driven entirely by noise uncertainty and was discussed by former literature (Mogstad et al., 2024; Andrews et al., 2024). Nevertheless, the expected mean regret from providing the personalized policy of one group to the other under uncertainty is similar to that when the decision-maker faces full certainty.

Policies based on the status quo average place more weight on the gains of natives and, therefore, provide higher regret for immigrants. This is illustrated in the red bars, which show that the status quo average policy nearly attains the outcomes of the personalized policy for natives. Under full information about location effects, the status quo average policy generates only 15.2 fewer shekels ($<$ \$5) for natives compared with the first-best. Likewise, the empirical Bayes policy attains 197 fewer shekels per year ($\approx$\$58), which is only 8% more than the personalized policy that ranks places by natives' posterior mean regret. Accordingly, the regret for immigrants from the status quo average policy is almost as bad as the personalized recommendation based only on the outcomes of natives. In the table right next to these figures, we show the share of selected places that are better than the status quo mean. Column 1 shows that even when θ is known, under the status-quo average policy, 4 out of the 10 selected cities will end up with regret higher than that expected under the status-quo sorting patterns for either group.

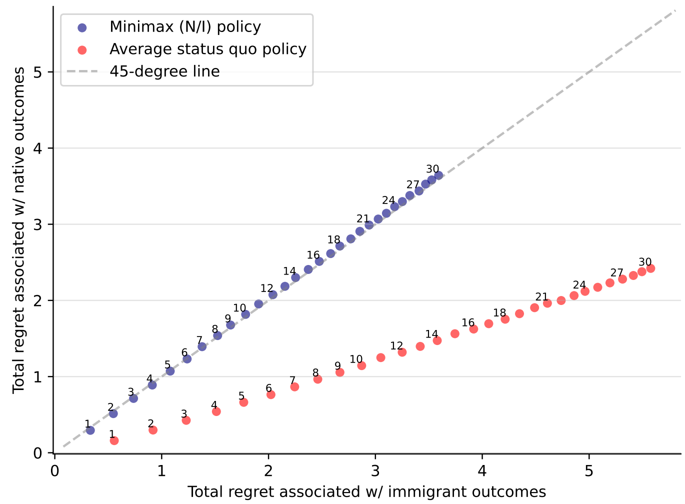
While the average status quo policy generates regret for minority groups, the purple and pink bars, which report the average regret under the minimax $(\mathcal{N}, \mathcal{I})$ and $(\mathcal{N}, \mathcal{I}$, city$)$ policies, show that it is possible to avoid extreme adverse outcomes. When ϑ is unknown (subfigure 6b), the minimax policies reduce immigrants' regret by a third compared to the average policy, at the cost of a 38% increase in regret for natives. The expected maximum regret across the immigration groups (the value of $\max(\mathcal{N}, \mathcal{I})$) drops by 15%, and the value of $\max(\mathcal{N}, \mathcal{I}, city)$—i.e., the worst-case scenario mentioned in Mogstad et al. (2024)—drops by 25%. Lastly, column 2 in the adjacent table shows that the minimax policies cannot provide full insurance against cities that are worse than the status quo, as in column 1. Still, they ensure that at most 2.5 of the 10 cities will generate regret higher than the average status quo value.

The gains from employing 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, Appendix Figure A.6a reports the expected regret from these policies for high-income families who exhibit a strong correlation between the location effects of immigrants and natives. For immigrants, the status quo average posterior mean policy provides only 54% higher regret than when places are ranked only based on the posterior mean of immigrants. This is much lower than the costs illustrated in Figure 6, for which the average status quo policy generates regret that is almost 3 times larger than the posterior mean of immigrants. As a result, minimax policies provide little improvement for immigrants compared to the average status quo policy. Worst case regret is only %15 lower, and the share of places that are worse than the status quo is only 6% higher than that under the average policy.

Fair policy: The trade-off faced by the decision-maker is visualized in Figure 7, which plots the expected total regret of immigrants and natives from the status-quo average and minimax ($\mathcal{N}/\mathcal{I}$) policies described in Equations (6) and (9). Each dot corresponds to the sum of selected cities' regret for K between 1 and 30, and the corresponding K is reported right next to the dot.

For every K, a policy based on the status quo average is clearly advantageous for natives but provides limited benefits for immigrants, which is reflected by the red curve lying below the 45-degree line. In contrast, for every K, the minimax strategy total regret lies very close to the 45-degree line, which implies a more equitable outcome. As K increases, total regret increases, both because we select more places and because it is harder to attain equal outcomes since fewer places 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 regret for natives, there exists a K for which the minimax strategy attains the same level of regret for natives with less regret for immigrants.

Figure 7: Total regret under minimax ($\mathcal{N}/\mathcal{I}$) and average status-quo targeting policies



Note: This figure plots the total regret of immigrants and natives from the average status-quo and minimax ($\mathcal{N}/\mathcal{I}$) policies. Blue dots plot outcomes generated by the minimax ($\mathcal{N}/\mathcal{I}$) policy described in Equation (9). Red dots plot outcomes generated by the status-quo average policy described in Equation (6). Curves are generated by varying K, the number of selected cities, between 1 and 30, and the K is printed next to each dot. Total regret is the sum of regret over selected cities, separately by immigration group, and the dashed line is the 45-degree line.

10.5 Selected Israeli Cities

Table 4 reports the top 15 Israeli cities sorted according to the posterior maximum regret across immigrants and natives when $K = 10$, where regret is in shekel values. The regret from selecting the top 10 leading cities is bounded for both groups. This is evident in columns 1-3, which report the posterior mean regret of immigrants, natives, and the city average, weighted by the observed status-quo shares. Compared with the first-best policy, an additional year spent in any of these cities generates, on average, 146-330 lower ILS earnings (43-97 US \$) for natives and lower 169-427 ILS earnings (50-125 US \$) for immigrants. Column (4) reports the posterior expectation of the within-city maximum regret ($\mathcal{N}, \mathcal{I}$). Even in the least favorable scenario, children in the top 10 cities have incomes only 237-470 ILS lower (equivalent to approximately 70-139 US \$) than in their average optimal 10 cities.

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 the regret of either immigrants or natives falls above each group's average under the status quo sorting, $\Pr(\vartheta_{j\mathcal{N}} \geq L_0 \text{ or } \vartheta_{j\mathcal{I}} \geq L_0 | \mathcal{Y})$, where $L_0 = \sum_{j=1}^{J} (\pi_{j\mathcal{I}0}\vartheta_{j\mathcal{I}} + \pi_{j\mathcal{N}0}\vartheta_{j\mathcal{N}})$, and π_{jg0} is the share of group $g \in \{\mathcal{N}, \mathcal{I}\}$ in city j in the status quo out of the entire population. This posterior probability is the analog of the false discovery rate in multiple hypothesis testing settings in which, for each city,

Table 4: Top Israeli cities selected based on minimax criterion, $K = 10$

	Posterior mean					
Loc. name	Native-born (1)	Imm. (2)	Average (3)	$E[\max\{\vartheta_{\mathcal{N}}, \vartheta_{\mathcal{I}}\} \| \mathcal{Y}]$ (4)	Worse than status-quo (5)	Selected by minimax ($\mathcal{I}/\mathcal{N}$/city) (6)
Bat Yam	145.8	168.3	161.2	237.1	0.011	Yes (680.2)
Haifa	222.9	293.6	274.4	344.2	0.090	Yes (680.2)
Rishon Leziyyon	268.9	270.7	270.4	350.6	0.073	Yes (680.2)
Holon	289.1	300.6	298.6	379.1	0.119	Yes (680.2)
Karmi'el	216.4	305.4	270.4	385.7	0.212	Yes (680.2)
Qiryat Gat	301.2	261.7	274.7	397.8	0.204	Yes (680.2)
Betar Illit	329.8	201.5	212.7	398.5	0.234	Yes (680.2)
Ashdod	67.2	418.4	289.8	421.9	0.274	Yes (680.2)
Dimona	233.8	412.2	371.0	468.2	0.368	Yes (680.2)
Arad	197.3	427.0	330.8	469.8	0.395	Yes (680.2)
Ma'alot-tarshiha	357.3	371.5	364.7	495.9	0.421	
Qarne Shomeron	294.0	404.9	387.7	497.5	0.432	
Efrata	433.1	240.9	267.6	499.4	0.441	
Yoqne'am Illit	291.6	430.1	393.5	506.6	0.443	
Qiryat Motzkin	465.7	257.3	303.1	507.2	0.451	

Note: This table reports a list of 15 Israeli cities sorted by within-city posterior immigrant-native maximum regret. Regret is the lost earnings at age 28 from spending one year in city j, compared with the average city selected under the first-best personalized policy. Columns 1-3 report posterior mean regret for natives, immigrants, and the average. Column 4 reports the posterior maximum regret across immigrants and natives. Column 5 gives the posterior probability that either group's location effect is below the average under status quo sorting. Column 6 reports the cities selected based on the minimax ($\mathcal{N}/\mathcal{I}/city$) policy, with the list's posterior maximum regret in parentheses.

we test the null that the regret of both immigrants and natives is greater than L_0. By averaging the first 10 values in column 5, we conclude that when selecting the top 10 cities following the minimax ($\mathcal{N}$,$\mathcal{I}$) policy, we should expect that at most 2 of these 10 cities would generate outcomes worse than the status quo for either of the two groups.

The last column indicates the cities selected under the minimax ($\mathcal{N}/\mathcal{I}/city$) policy, which, in our data, coincides with the list of cities selected by the minimax ($\mathcal{N}$,$\mathcal{I}$) policy when $K = 10$. In parentheses, we report the posterior expectation of the maximum regret across all selected cities and across both immigrants and natives. 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. It provides us with a guarantee that, on average, regret would not be higher than 680 ILS.

10.6 Model Extension and Robustness

Model Extension: The stylized models depicted in Section 10.3 provide a clear, easy-to-interpret closed-form decision rule for a minimax decision-maker who seeks robustness against the least favorable sorting scenario. Nevertheless, these policies are derived under simplified assumptions that may not hold in reality. First, the minimax decision-maker behaves as if all families might sort into a single worst place—a phenomenon that is rejected by the data. To illustrate this, Appendix Figure A.7 plots the distribution of location choices of families if they follow the minimax ($\mathcal{N}/\mathcal{I}/city$) strategy and face the minimax ($\mathcal{N}/\mathcal{I}/city$) decision-maker. This figure shows that the minimax behavior assumes that families sort into a small set of places, which might not be reasonable. Second, the models in Section 10.3 do not take into account capacity constraints or other limitations that may arise in real-world settings.

To account for these concerns, Appendix Section L describes an extended model that restricts sorting probabilities to better align with the spatial distribution in the data while maintaining compliance uncertainty. We show that when location choices are restricted to align closely with the status quo spatial distribution, decisions are similar to those of the average status quo policy. In contrast, with more ambiguity regarding location choices, optimal policy aligns with the minimax decision rule and offers more equal outcomes to both groups. On that scale of ambiguity, the choice between these possible models depends on the decisionmaker's information and uncertainty.

The range of restrictions on location choices discussed in this section reflects the importance of careful contemplation of the information set and social objectives decision-makers might have. While previous literature has primarily focused on how noise affects decisions (Mogstad et al., 2024; Andrews et al., 2024), this extension emphasizes that statistical uncertainty alone could result in various types of policies, depending on the uncertainty decision-makers face on other dimensions.

Normalization: To assess the sensitivity of the results to the regret normalization, Appendix Table A.7 replicates Table 4 while normalizing the value of each place in comparison with that expected under the status quo sorting patterns. Selecting the top 10 Israeli cities using the mean status quo normalization yields the same list as in Table 4, although the within-list ranking is different.

Noisy hyperparameters: Appendix section G.3 studies the sensitivity of our analysis to noisy hyperparameters. Using a full hierarchical Bayesian approach, we estimate the model with Hamiltonian Monte Carlo (HMC), setting a flat prior on the mean and variance hyperparameters. The resulting hyperparameter estimates and list of selected cities align closely with the ones estimated in our 2-step EB procedure.

11 Conclusion

This paper studies the heterogeneity in the causal effects of Israeli cities on children's income in adulthood. Our exploration into the nuanced differences between childhood location effects of natives vs. immigrants in Israel has shown that cities benefiting one group do not necessarily benefit the other.

Previous literature has used similar city-wide average location effects to rank places and inform recommendations for families relocating through housing voucher programs (Bergman et al., 2019). While the literature on school value-added and hiring policies has emphasized the importance of taking into account match effects and treatment effect heterogeneity (e.g., Biasi et al., 2021; Bates et al., 2024), neighborhood recommendation policies often treat places as an ordered treatment, usually proportional to mean poverty rate or mean earnings (Katz et al., 2001; Bergman et al., 2019). Our findings suggest that such policies may disproportionately harm minorities.

We discuss the trade-offs policymakers face when implementing a unified policy that cannot be conditioned on individual characteristics, in settings with both noise and compliance uncertainty. While the literature has primarily focused on the risks of forming policy based on noisy estimates, we highlight that uncertainty driven by heterogeneity should also be taken into account when treatment effects vary substantially. Nevertheless, using a decision-theoretic framework, we show that by acknowledging the ambiguity with respect to individuals' sorting behavior, it is possible to find at least 10 cities in Israel that are beneficial to both groups. Our model demonstrates that fairness can be improved even in restricted unified policies. Such a model could be useful in many other settings where treatments are directed not at individuals but at predefined groups. While we illustrate the advantage of our model in the context of immigrants and natives in Israel, the model generalizes, as heterogeneity in place effects has been documented in other contexts (Chetty et al., 2018, 2020).

References

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.

Abramitzky, R., Boustan, L., and Connor, D. S. (2024). Leaving the enclave: Historical evidence on immigrant mobility from the industrial removal office. *The Journal of*

Economic History, 84(2):352–394.

Abramitzky, R., Boustan, L., Jácome, E., and Pérez, S. (2021). Intergenerational mobility of immigrants in the united states over two centuries. *American Economic Review*, 111(2):580–608.

Aloni, T. (2017). Intergenerational mobility in israel. *School of Economics, Tel-Aviv University*, M.A. Dissertation.

Andrews, I., Kitagawa, T., and McCloskey, A. (2024). Inference on winners. *The Quarterly Journal of Economics*, 139(1):305–358.

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).

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.

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.

Bates, M., Dinerstein, M., Johnston, A. C., and Sorkin, I. (2024). Teacher labor market policy and the theory of the second best. *The Quarterly Journal of Economics*.

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.

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.

Biasi, B., Fu, C., and Stromme, J. (2021). Equilibrium in the market for public school teachers: District wage strategies and teacher comparative advantage.

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., 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.

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.

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., Jones, M. R., and Porter, S. R. (2020). Race and economic opportunity in the united states: An intergenerational perspective. *The Quarterly Journal of Economics*, 135(2):711–783.

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.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., et al. (2022). Social capital i: measurement and associations with economic mobility. *Nature*, 608(7921):108–121.

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. (2022). The long-run effects of residential racial desegregation programs: Evidence from gautreaux.".

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

Cohen, S., Hsieh, C.-T., et al. (2001). *Macroeconomic and labor market impact of Russian immigration in Israel*. Bar-Ilan Univ.

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.

Derenoncourt, E. (2022). Can you move to opportunity? evidence from the great migration. *American Economic Review*, 112(2):369–408.

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? *Unpublished draft*.

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

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

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., Ludwig, J., Mullainathan, S., and Rambachan, A. (2018). Algorithmic fairness. *AEA papers and proceedings*, 108:22–27.

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. (2024). A discrimination report card. *American Economic Review*, 114(8):2472–2525.

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

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.

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

Mogstad, M., Romano, J. P., Shaikh, A. M., and Wilhelm, D. (2024). Inference for ranks with applications to mobility across neighbourhoods and academic achievement across countries. *Review of Economic Studies*, 91(1):476–518.

Rambachan, A., Kleinberg, J., Mullainathan, S., and Ludwig, J. (2020). An economic approach to regulating algorithms.

Savage, L. J. (1951). The theory of statistical decision. *Journal of the American Statistical association*, 46(253):55–67.

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

Online Appendix

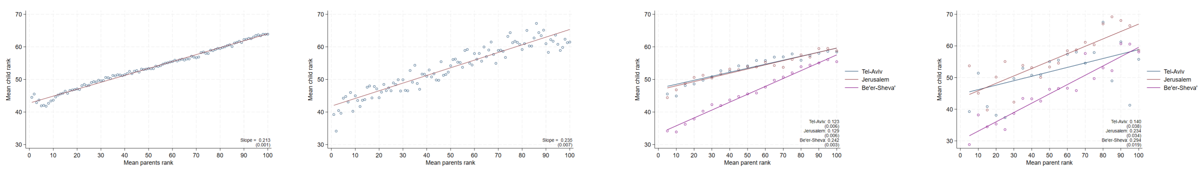
(Not For Publication)

A Additional Figures and Tables

Figure A.1: Parental income rank and child mean income rank at ages 28–30

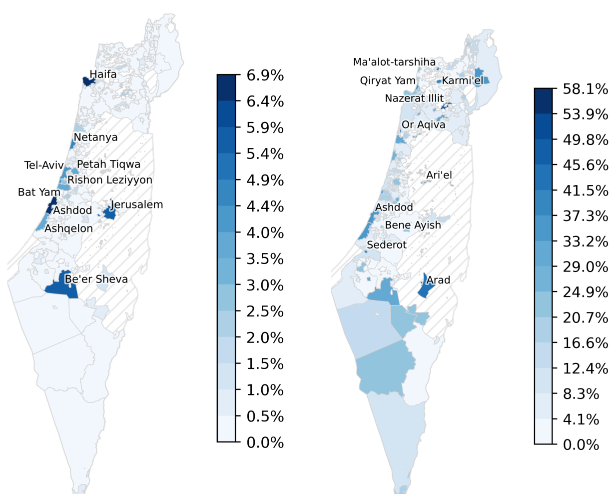
All Cities | **Selected Cities**

(a) Natives (b) Immigrants (c) Natives (d) Immigrants



Note: These figures display the relationship between parental income rank and mean child income rank for immigrants from the Former Soviet Union and native-born Israeli children in ages 28–30. Left panels show all Israeli cities; right panels focus on children who lived in Tel Aviv, Be'er Sheva, and Jerusalem from birth to age 18.

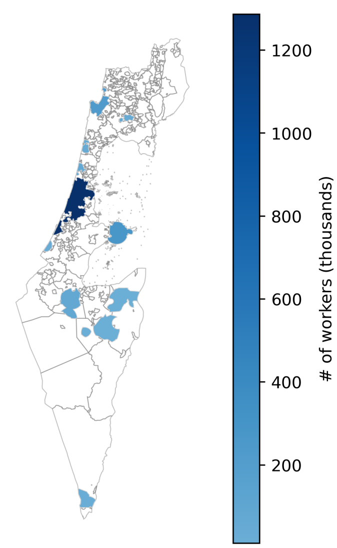
Figure A.2: Immigrants' spatial distribution



(a) Across city share (b) Within city share

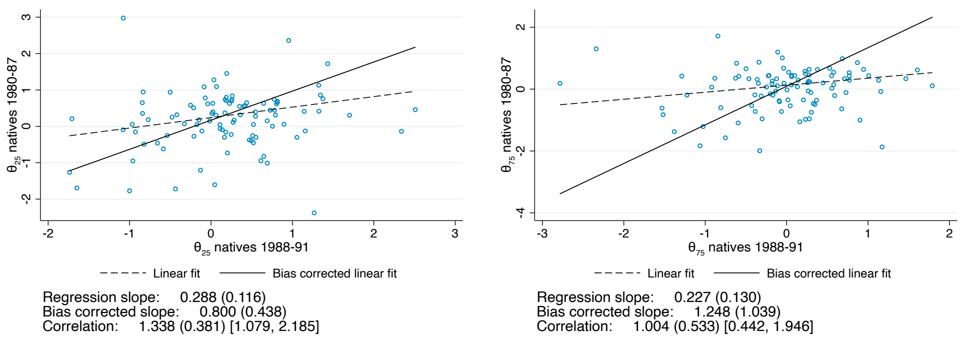
Note: This map presents the geographic distribution of immigrants across Israel. Panel (a) maps the share of immigrants in each location out of the immigrant population. Panel (b) maps the share of immigrants out of the whole population within each city. Location names are attached to the cities with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

Figure A.3: Employment centers



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

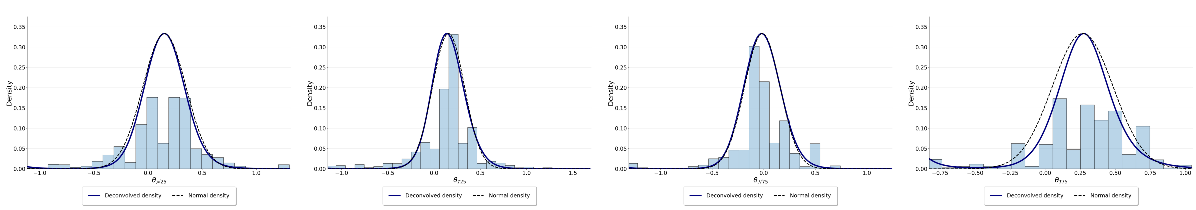
Figure A.4: The relationship between location effects of natives born in 1980-1987 and 1988-1991



(a) Total effect, θ_{25} (b) Total effect, θ_{75}

Note: These figures display the scatter plots and observation-weighted regression lines of location effects for native Israeli children born in years 1980-1987 and 1988-1991. Panel (a) presents the location effects for families in the 25th percentile of the income distribution, and panel (d) the location effects for families in the 75th percentile of the income distribution. The dashed line is the naive regression line and the solid line is the bias-corrected regression. Square brackets display parametric bootstrapped equal-tailed confidence intervals

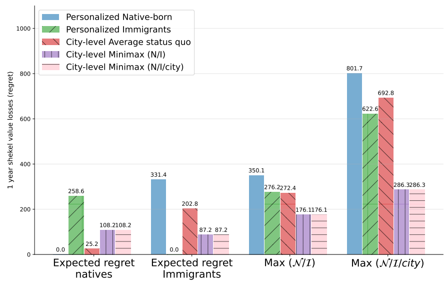
Figure A.5: Deconvolved density of immigrants' and natives' location effects (θ)



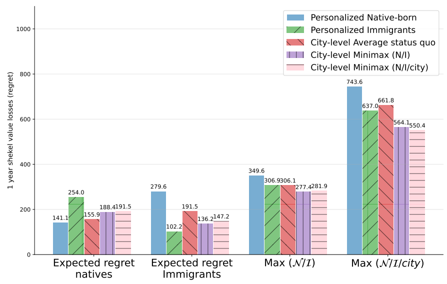
(a) θ_{25} Natives (b) θ_{25} immigrants (c) θ_{75} Natives (d) θ_{75} immigrants

Note: These figures plot the distribution of the location effects of immigrants and natives estimated in Equation (2). Panels (a) and (b) present the distribution for low-income natives and immigrants ($p = 25$), and panels (c) and (d) present the distribution for high-income natives and immigrants ($p = 75$). The solid blue line shows the weighted Efron (2016) deconvolved log-spline density, and the histograms show the weighted estimated location effects, both weighted by city size in 2003. Dashed lines show the density of a normal distribution with the same mean and variance.

Figure A.6: Targeting trade-offs from choosing the top 10 cities, $p = 75$



(a) Known θ (oracle)



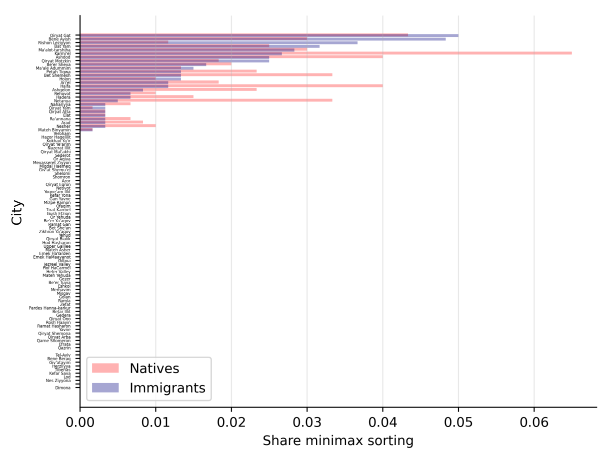
(b) Unknown θ

	% better than the status-quo	
	Known θ (oracle) (1)	Unknown θ (2)
i) Personalized policy		
Native-born	0.742	0.623
Immigrants	0.585	0.703
ii) City-level policy		
Av. status quo	0.751	0.711
Minimax ($\mathcal{N}/\mathcal{I}$)	1.000	0.773
Minimax ($\mathcal{N}/\mathcal{I}$/city)	1.000	0.762

(c) Share better than the status-quo

Note: These plots present the expected regret from different policies that aim to select the top 10 places in Israel based on the location effects for children whose parents are in the 75th percentile of the income distribution. Subfigure (a) displays results from policies based on the true location effects, and subfigure (b) displays the results from policies based on the empirical Bayes posterior mean values. The first two groups of bars report the expected regret of each policy and group. The $Max(\mathcal{N}, \mathcal{I})$ bars report the expected within-city immigrant-natives maximum regret of selected cities, and the $Max(\mathcal{N}, \mathcal{I}, city)$ bars report the expected maximum regret across all selected cities and immigrants and natives. The blue and green bars report the results from a policy that ranks locations based on the regret of each group. The red bars report the results from a policy that ranks locations based on the status-quo average regret described in Equation (6). The purple bars report the results from the minimax ($\mathcal{N}/\mathcal{I}$) policy described in Equation (9). The pink bars report the results from the minimax ($\mathcal{N}/\mathcal{I}$/city) policy described in Equation (12). Subplot (c) presents the share of selected cities with regret lower than that under the status quo sorting patterns of each group.

Figure A.7: 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 decision-maker. Red bars display the sorting probabilities of native families, and blue bars display the sorting probabilities of immigrant families.

Table A.1: Variance component of the intercept and slope $(\alpha_{j\mathcal{N}}, \eta_{j\mathcal{N}}, \alpha_{j\mathcal{I}}, \eta_{j\mathcal{I}})$

	Natives		Immigrants	
	Cons. (1)	Rank-parents (2)	Cons. (3)	Rank-parents (4)
Natives				
Cons.	0.245			
	(0.042)			
Rank-parents	-0.724	0.003		
	(0.124)	(0.000)		
	[-0.880, -0.372]			
Immigrants				
Cons.	0.041	-0.166	0.194	
	(0.278)	(0.271)	(0.049)	
	[-0.441, 0.538]	[-0.668, 0.297]		
Rank-parents	-0.215	0.708	-0.457	0.003
	(0.223)	(0.188)	(0.177)	(0.000)
	[-0.676, 0.228]	[0.369, 1.030]	[-0.690, 0.052]	

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	0.200	0.128	0.360
	(0.037)	(0.059)	
Std. Immigrants	0.173	0.101	0.416
	(0.043)	(0.052)	
Std. Difference	0.266	0.545	
	(0.055)	(0.154)	
Immigrants-native corr	-0.017	-0.154	
	(0.320)	(0.673)	
	[-0.500, 0.444]	[-1.101, 0.711]	
(ii) High-income families (θ_{75})			
Std. Natives	0.172	0.097	0.436
	(0.037)	(0.036)	
Std. Immigrants	0.218	0.179	0.188
	(0.032)	(0.057)	
Std. Difference	0.209	0.473	
	(0.053)	(0.166)	
Immigrants-native corr	0.445	-0.076	
	(0.239)	(0.476)	
	[0.010, 0.901]	[-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 component, equal neighborhood weights

	Low-income families, θ_{25}		High-income families, θ_{75}	
	Std. (1)	χ^2 test $H_0: \theta_{jI} - \theta_{jN} = c\ \forall j$ (2)	Std. (3)	χ^2 test $H_0: \theta_{jI} - \theta_{jN} = c\ \forall j$ (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 the location effects of native-born and immigrants. 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) report the standard deviation of location effects, while the third row reports the correlation between natives' and immigrants' location effects. Standard errors are calculated via delta method.

Table A.4: Heterogeneity in location effects, within group parents income rank

	Low-income families, θ_{25}		High-income families, θ_{75}	
	Natives (1)	Immigrants (2)	Natives (3)	Immigrants (4)
Natives	0.158		0.193	
	(0.052)		(0.041)	
Immigrants	-0.116	0.186	0.488	0.248
	(0.338)	(0.026)	(0.215)	(0.025)
	[-0.689, 0.439]		[0.064, 0.882]	

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.

Table A.5: Heterogeneity in location effects, earnings and log earnings

	Low-income families, θ_{25}		High-income families, θ_{75}	
	Natives (1)	Immigrants (2)	Natives (3)	Immigrants (4)
A) Earnings				
Natives	169.53		262.42	
	(300.71)		380.80)	
Immigrants	-0.75	367.35	0.35	557.07
	(1.48)	(61.53)	(1.18)	(59.98)
	[-3.12, 0.96]		[-0.81, 1.16]	
B) Log earnings				
Natives	0.0181		0.0179	
	(0.0081)		(0.0068)	
Immigrants	0.0284	0.0268	0.5638	0.0350
	(0.4008)	(0.0042)	(0.3220)	(0.0043)
	[-0.614, 0.324]		[0.122, 1.138]	

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 \$ $\approx$ 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.

Table A.6: Heterogeneity in location effects, 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 deviations and correlation of the location effects of immigrants and natives for different reweighting schemes. Columns 1-2 report the correlation matrix for low-income families, and columns 3-4 report the correlation matrix for high-income families. 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 via delta method. Square brackets display parametric bootstrapped equal-tailed confidence intervals.

Table A.7: Top selected Israeli cities, $K = 10$, status-quo sorting normalization

	Posterior mean					
Loc. name	Native-born (1)	Imm. (2)	Average (3)	$E[\min\{\vartheta_{\mathcal{N}}, \vartheta_{\mathcal{I}}\}\|\mathcal{Y}]$ (4)	Worse than status-quo (5)	Selected by minimax ($\mathcal{I}/\mathcal{N}$/city) (6)
Bat Yam	184.4	347.0	295.4	161.0	0.087	Yes (-234.6)
Ashdod	263.0	97.0	157.7	72.7	0.295	Yes (-234.6)
Haifa	107.2	221.7	190.6	70.7	0.245	Yes (-234.6)
Rishon Leziyyon	61.2	244.6	210.3	41.0	0.359	Yes (-234.6)
Karmi'el	113.7	209.9	172.1	36.0	0.396	Yes (-234.6)
Holon	41.0	214.7	184.0	16.6	0.437	Yes (-234.6)
Qiryat Gat	28.9	253.6	180.1	-7.3	0.512	Yes (-234.6)
Arad	132.8	88.3	107.0	-16.5	0.518	Yes (-234.6)
Betar Illit	0.4	313.8	286.5	-24.2	0.539	Yes (-234.6)
Dimona	96.4	103.1	101.6	-26.0	0.544	Yes (-234.6)
Ashqelon	291.7	-41.0	75.8	-52.9	0.578	
Qarne Shomeron	36.1	110.4	98.9	-71.1	0.633	
Yoqne'am Illit	38.5	85.2	72.9	-74.1	0.645	
Ma'alot-tarshiha	-27.2	143.8	62.0	-90.5	0.691	
Gush Etzion	-16.8	121.5	101.6	-95.2	0.683	

Note: This table reports a list of 15 Israeli cities sorted by within-city posterior immigrant-native minimum effect. Columns 1-3 report the posterior mean of native-born children, immigrants, and the average. Column 4 reports the posterior minimum between immigrants and natives. Column 5 reports the posterior probability that the location effects of immigrants or the location effects of natives are lower than their average effect under the status quo sorting patterns. Column 6 reports the cities selected based on the minimum ($\mathcal{N}/\mathcal{I}/city$) policy, where the posterior maximum effect of the selected list is presented in parentheses.

B Data and Definitions

The data covers the entire Israeli population born between 1950 and 1995, including birth/death years, and matched parents. We merge this with the 1995-2019 tax records and the annual population registry files (1995, 1999–2019), which report city and statistical area of residence for children and parents, as well as immigration year and origin country. Earnings are inflation-adjusted to 2016 prices, summed across all sources (employed and self-employed). Income ranks are estimated on the entire population. A child is an immigrant if at least one parent was born in the USSR and immigrated between 1989-2000. We drop anomalous records: birth after death (≈ 450), parent birth year before 1950 (10), unmatched parent IDs in tax data (113), and negative earnings (≈ 950).

Using registry data as the main source for locality, we construct geographic mobility variables. Since we observe each parent's and the child's location annually, we define a child's location based on the parent with the most overlapping years in the same locality. A move is recorded when the locality changes from one year to the next. For the missing registry years (1996–1998), we use school locality data to identify moves when it matches either the origin or destination. We further enrich location data using the 1995 census, which includes responses to "When did you move to your current city?" and "Where did you live 5 years ago?"—allowing us to extend mobility histories back to 1990 for some individuals. Sample is restricted to localities with at least 100 children of each group, immigrants and natives. This yields 98 of the 256 localities and regional councils.

City Level Covariates:

- **Gini index (1998)**: City-level inequality, based on gross earnings from National Insurance Institute data, reported in the Local Authorities in Israel report.
- **Diversity**: is the within-city group entropy: $E_d = \sum_g \pi_{g\in\{\mathcal{N},\mathcal{I}\}} \log \frac{1}{\pi_g}$, where π_g is the share of group g individuals in the city, and the summation is performed over all the groups in the city. E_d ranges from 0 (homogeneous) to $\log G$ (maximum diversity with G groups).
- **Theil (1972) index for segregation:** constructed using the 2000 earnings records. For every group g (immigrants or natives), we compute the city-level entropy E and sub-area entropy E_s. Segregation is: $H = \sum_s (\frac{pop_s}{pop} \frac{E-E_s}{E})$, where pop_s is the population size of subarea s and pop that city population size.
- **Share of criminal offenders (2002)**: Share charged with serious crimes.
- **Municipality welfare expenditure**: Municipality-level spending on welfare per capita (1998), from administrative data via the Local Authorities report.
- **High-school Bagrut eligibility**: Municipality-level share of 12th graders eligible for a Bagrut

certificate in 1999–2000, based on Ministry of Education records.

- **Distance to employment center**: The distance from each city to the nearest employment hub, based on the Central Bureau of Statistics 2008 census workplace data.
- **Peripherality index**: Developed by the Central Bureau of Statistics, measuring cities geographic proximity to major population centers. It is based on factors like distance from markets, employment hubs, and the Tel Aviv district.

C One Move vs. Multiple Moves

Tables C.1 and C.2 describe the sample of immigrants and natives by number of moves, and Tables C.3 and C.4 replicate the main analysis in the paper in the sample to the single move population.

Table C.1: Descriptive statistics, immigrants

	All cities			98 sample cities		
	Stayers (1)	Movers (2)	All (3)	Stayers (4)	Movers (5)	All (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
(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
Number 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), by number of moves between cities in Israel. Stayers refer to migrants who were either born in Israel or immigrated and did not move within Israel by age 18, and movers refer to those who moved between Israeli cities once. The left panel includes all cities in our sample, and the right includes only the 98 cities for which we can estimate the effects for both immigration groups. Income is measured in Israeli Shekels (≈ 3.4$).

Table C.2: Descriptive statistics, natives

	All cities			98 sample cities		
	1 move (1)	2 moves (2)	Stayers (3)	1 move (4)	2 moves (5)	Stayers (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
(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
Number of children	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), by number of moves between cities in Israel. The left panel includes all cities in our sample, and the right includes only the 98 cities for which we can estimate the effects for both immigration groups. Income is measured in Israeli Shekels (≈ 3.4$).

D Research Design Validation

Balance Test: Identification assumption requires that exposure time is 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. The figures below provide our test for these assumptions.

Figure D.1 shows the relationship between native-born children's age at first move and parental characteristics. Sub-figures D.1b and D.1a display raw associations (gray dots

Table C.3: Heterogeneity of location effects, 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 α and η							
Natives							
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							
Natives							
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 reports estimates of the distribution of location effects for immigrant and native children. Columns (1)-(3) report estimates for all cities, while columns (4)-(6) report estimates for cities with sufficient samples for both immigrants and natives. Columns (2) and (5) report the mean, and columns (3) and (6) the square root of the bias-corrected variance component. Standard errors for all variance estimators are based on the asymptotic variance, assuming location effects are drawn from a normal distribution.

Table C.4: Differences in location effects between immigrants and natives, single move

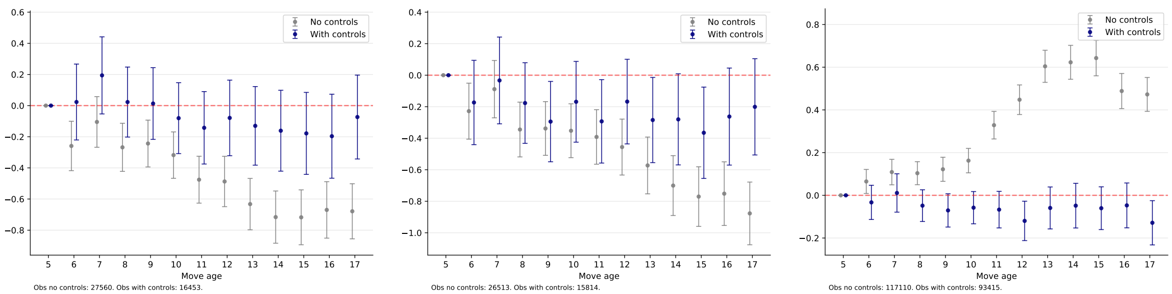
				Difference		
	Covariance	Correlation	Implied OLS coefficient	Mean	Std.	χ^2 test $H_0 : \theta_{jI} - \theta_{jL} = c \; \forall j$
	(1)	(2)	(3)	(4)	(5)	(6)
α	-0.0047	-0.0839	-0.069	0.5189	0.3571	156.6
	(0.0180)	(0.3213)	(0.266)	(0.1346)	(0.0797)	[0.0000]
		[-0.6407, 0.4661]	[-0.505, 0.335]			
η	0.0000	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 natives. Column (1) presents the covariance, column (2) presents the bias-corrected correlation, which is the covariance divided by the product of standard deviations of immigrants and natives, 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 χ^2 test for the null that location effects don't vary within cities. 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.

and error bars) and covariate-adjusted ones (blue dots), controlling for the variables listed in Section 3.2. Parents' years of schooling are obtained from the 1995 census, which is available for 20% of the population. The more educated the parents, the more likely they are to move when their children are younger. However, this relationship disappears after controlling for x_i, suggesting no systematic relationship with time-invariant characteristics.

Sub-figure D.1c shows the relationship between age at move and parents' earnings growth during ages 0–5, using salary-employed earnings from 1986–1995, which are not used to construct parental income ranks.[35] The raw relationship suggests that families who moved when children were older had higher early earnings growth. However, after controlling for x, this relationship balances, supporting the assumption that time-varying factors are not systematically related to age at move after controlling for x_i. Similarly, Figure D.2 shows the relationship between children's age at arrival in Israel and parental schooling. Figures D.4 and D.3 repeat the analysis for the child's age at the second move.

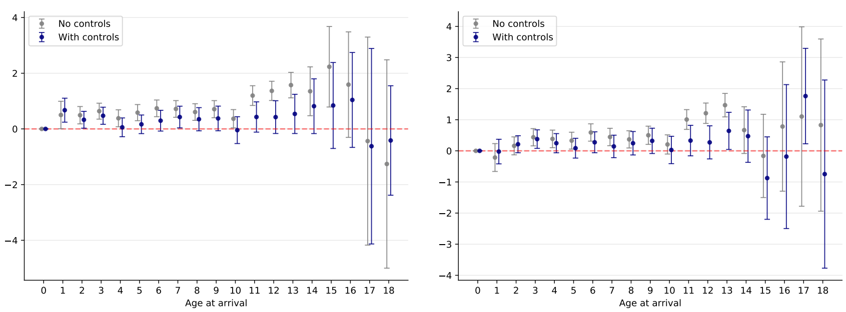
Figure D.1: Age at first move and parents' characteristics, native-born children



(a) Mother years of schooling (b) Father years of schooling (c) Wage growth, ages 0-5

Note: This figure presents the relationship between the child's age at the move and the parents' education. Confidence intervals are based on family-level clustered standard errors.

Figure D.2: Age of arrival to Israel and parents' characteristics, immigrants



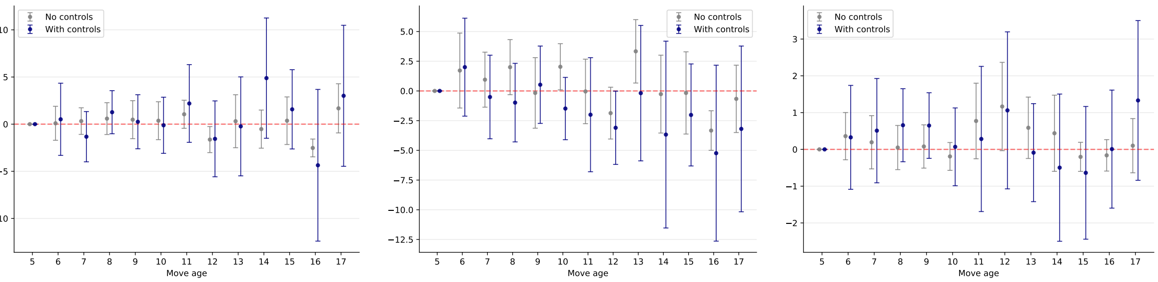
(a) Father years of schooling (b) Mother years of schooling

Note: This figure presents the relationship between the child's age at the move to Israel and the parents' education. Confidence intervals are based on family-level clustered standard errors.

Moving to cities with higher mean outcomes: To test the linearity assumption and identify the age at which childhood location no longer impacts children's outcomes, we use a split-sample approach to test whether moving earlier to a city with higher average outcomes improves children's long-run outcomes (Chetty and Hendren, 2018). For natives, we estimate each city's mean age 28 income rank, $\bar{Y}_j$, using children whose families never

[35] It is excluded due to the missing self-employment income.

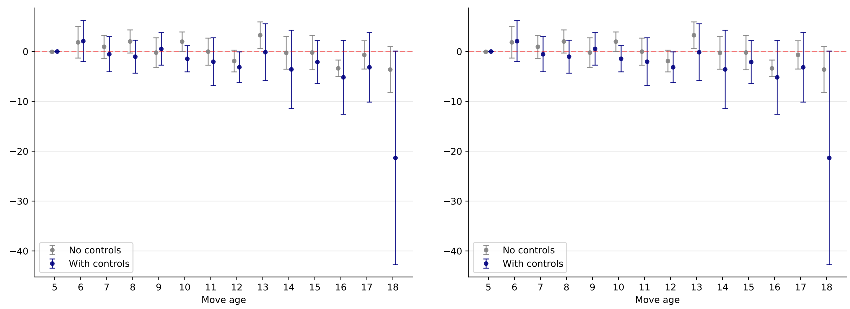
Figure D.3: Age at second move and parents' characteristics, native-born children



(a) Mother years of schooling (b) Father years of schooling (c) Wage growth ages 1-5

Note: This figure presents the relationship between the child's age at the second move and the parents' education and earnings. Confidence intervals are based on family-level clustered standard errors.

Figure D.4: Age at second moved between 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 second move to Israel and the parents' education. Controls include origin-destination-birth-year fixed effects and parents' income rank interacted with the year of birth. Confidence intervals are based on family-level clustered standard errors.

moved between cities before age 30. Then, in the sample of native-born children whose families moved only once, we estimate the following regression:

$$Y_i = \sum_{m=1}^{30} \beta_m \mathbb{1}\{m(i) = m\} \Delta_{o(i)d(i)} + x_i'\gamma + \epsilon_i, \tag{13}$$

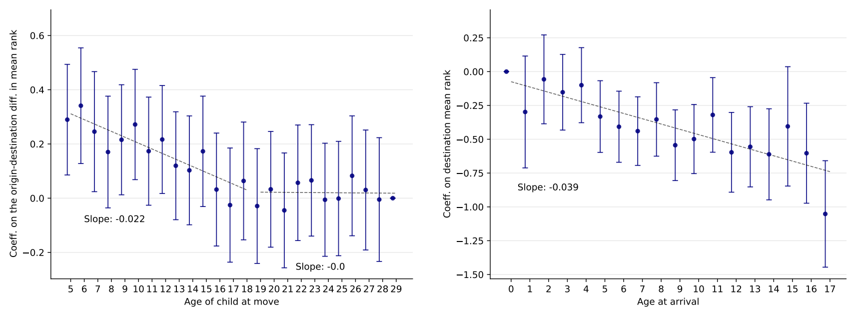
where $m(i)$ is child's i move age, $o(i)$ and $d(i)$ are child's i origin and destination cities, x_i includes the controls described in section 3.2, and $\Delta_{od} = \bar{Y}_d - \bar{Y}_o.\beta_m$ are the parameters of interest, which measure the effect of moving at age m to a city with one percent higher mean income rank. For immigrants, we compute $\bar{Y}_j$ with the Israeli-born immigrant children who remained in the same city until age 17. Since they are included in the main sample, we randomly split the Israeli-born immigrant sample into two subsamples, $s \in 1, 2$, and compute $\bar{Y}_{js}$ within each. Then, on the sample of immigrants who stayed in the same Israeli city until age 18, we estimate:

$$Y_i = \beta_0 \mathbb{1}\{a(i) = 0\}\bar{Y}_{j(i)s'(i)} + \sum_{a=1}^{17} \beta_a \mathbb{1}\{a(i) = a\}\bar{Y}_{j(i)} + x_i'\gamma + \epsilon_i, \tag{14}$$

where $s'(i)$ is the split that excludes child i, and β_a are the parameters of interest, which measure the effect of arriving in Israel at age a to a city where Israeli-born immigrants have a one percentile higher average income.

Figure D.5 plots the results. The estimates for natives show that β_m declines linearly for ages $m < 18$, indicating that earlier moves to higher-opportunity cities are associated with higher income ranks at age 28–a pattern consistent with findings from other countries. For ages $m \geq 18$, the effect remains flat, consistent with the Israeli context, where most children enlist in the army at 18–19 and are thus less exposed to residential location. Figure D.5 shows a similar pattern for immigrants.

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 at age 28 for natives (a) and immigrants (b). Confidence intervals are based on family-level clustered standard errors.

The kink at age 18 motivates a piece-wise regression, reported in columns (1) and (3) of Table D.1. For natives, the slope above age 18 is small and statistically insignificant. The slope below age 18 is 0.022 for natives and 0.033 for immigrants. These effects are in line with, though slightly smaller than, previous findings in the U.S. (0.035 at the county level, Chetty and Hendren, 2018), Australia (0.033, Deutscher, 2020), and Canada (0.042, Laliberté, 2021).

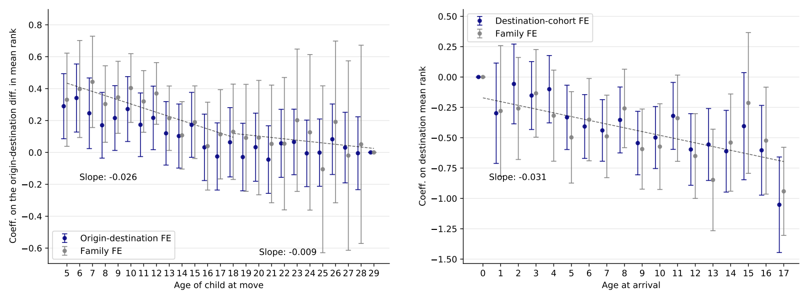
Robustness to family fixed effects: Figure D.6 presents the estimates from Equations 13 and 14, with and without family fixed effects. Blue dots replicate the estimates from Figure D.5, while gray dots show estimates controlling for family FEs. The linear decline in β_m (natives) and β_a (immigrants) up to age 18 remains consistent, though noisier. Similarly, columns (2) and (4) in Table D.1 show family FE slopes of 0.019 for natives and 0.026 for immigrants, similar to those in columns (1) and (3).

Table D.1: Years of exposure in childhood city and mean outcomes

	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 of the exposure model. Columns (1) and (3) include the controls described in Section 3.2, and columns (2) and (4) include family fixed effects. Columns (5) and (6) present the slope coefficients on the posterior means and the arrival age for immigrants and the age of move for natives. Standard errors in parentheses are clustered at the family level.

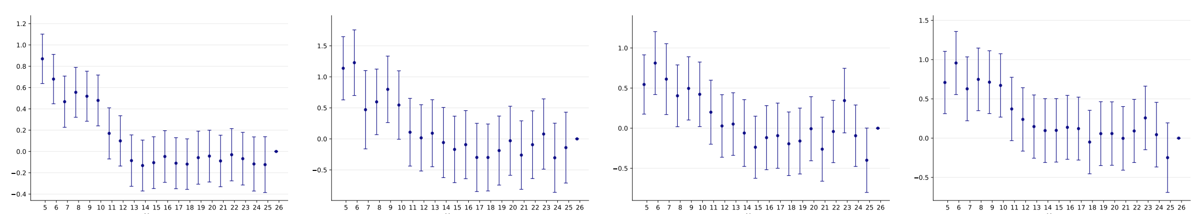
Figure D.6: Robustness to family fixed effects



(a) Natives (b) immigrants

Note: This figure presents the exposure effect coefficients on child's earnings rank at age 28 for natives (a) and immigrants (b). Dashed lines are the linear fit with the corresponding slopes written below. Panel (a) presents the effects for native-born Israeli children, and Panel (b) for immigrant children. Gray dots show results from a model with family fixed effects. Blue dots are the results from Figure D.5. Confidence intervals are based on standard errors clustered by family id.

Figure D.7: Exposure effects on 5th grade standardized exam scores for natives



(a) Math (b) Science (c) English (d) Hebrew

Note: This figure presents the exposure effect coefficients for the 5th grade *Meitzav* standardized exam score. Different panels display results for students' scores in different subjects: mathematics (a), science (b), English (c), and Hebrew (d), measured in standard deviations. Confidence intervals are based on family-level clustered standard errors.

Outcomes realized in childhood: Figure D.7 presents a placebo test using outcomes realized before age 18, by estimating Equation 13 on children's scores in the 5th-grade national standardized exams, *Meitzav*. These exams cover subjects such as math, science, English, and Hebrew/Arabic, resembling the OECD's PISA test. They are administered to a representative sample of schools, covering about 12% of the natives. Since the exam is taken at age 11, estimates from ages 11–18 serve as a placebo test. Note that because children who immigrated after the age of 11 never took the exam, we can perform this exercise only for natives.

Test for linearity - The city-level mean outcomes used above combine location effects and average ability, $\bar{Y}_j = \theta_j + \bar{\xi}_j$. Thus, the relationship in Figure D.5 could reflect either a linear relationship with location effects or a mix of linear relationships with effects and with $\bar{\xi}_j$. To test whether the true location effects θjg themselves linearly relate to years of exposure, we conduct an additional exercise. We regress the estimated location effects of movers from Equation (2) on children's income rank by child's age at move. To avoid mechanical correlation, we run Equation (2) on two random split samples $s(i) \in \{1, 2\}$. Then, for every group g, city j, and sample split s, we estimate $\hat{\theta}^s_{jgp}$ and use it as a covariate in the following model for natives:

$$Y_i = \sum_{m=1}^{18} \beta_m \mathbb{1}\{m(i) = m\} \hat{\Delta}^{s'(i)}_{o(i)d(i)p(i)} + x_i'\gamma + \epsilon_i \tag{15}$$

where $s'(i)$ is the random sample that excludes child i, $\hat{\Delta}^s_{od} = \hat{\theta}^{*s}_{dgp} - \hat{\theta}^{*s}_{ogp}$, and $\hat{\theta}^{*s}_{dgp}$ are the Empirical Bayes posterior means of θ_{dgp}. Similarly, for immigrants, we estimate :

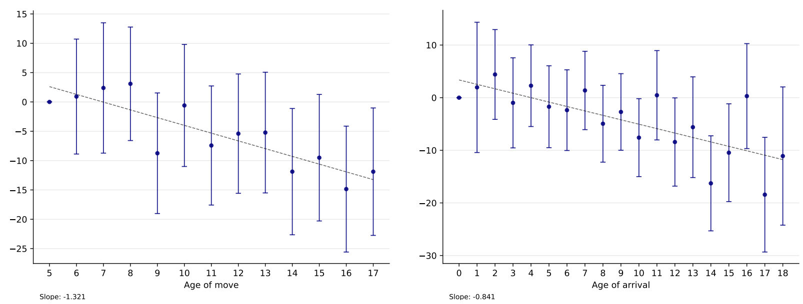
$$Y_i = \sum_{a=1}^{18} \beta_a \mathbb{1}\{a(i) = a\} \hat{\theta}^{*s}_{j(i)g(i)} + x_i'\gamma + \epsilon_i \tag{16}$$

where Israeli-born immigrants are the base level. If exposure effects are linear, then for immigrants, the coefficient β_a gives us the *number* of θ_j's accumulated in location j as θ^* is the one-year effect of city j. Thereby, the slope, $\beta_{a+1} - \beta_a$, with respect to exposure time, should be 1 on average. The results, presented in Figure D.8 and Table D.1, show a linear relationship between the posterior means and the age of move, with fitted slopes of 1.321 and 0.841, statistically indistinguishable from 1.

E Variance Componenets

Denote $\Omega \in \mathbb{R}^{4\times 4}$ 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 is $\sigma_g^2 = \sum_{j=1}^{J} \pi_j (w_{jg} - \sum_{j=1}^{J} \pi_j w_j)^2$, where $w_{jg} \in \{\alpha_{jg}, \eta_{jg}\}$, and $\pi_j = \frac{n_j}{N}$ are city shares , with n_j is city j population size, and

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



(a) Native-born (b) Immigrants

Note: This figure displays the β_a coefficients from Equations (15) for immigrants and (16) for natives. Confidence intervals are based on standard errors clustered by family id.

$N = \sum_{j=1}^{J} n_j$. Note that we can write the variance also as:

$$\sigma_g^2 = \sum_{j=1}^{J}(1-\pi_j)\pi_j w_{jg}^2 - 2\sum_{j=1}^{J}\sum_{k=j+1}^{J} \pi_j \pi_k w_{jg} w_{kg} \equiv S_{ML}. \tag{17}$$

Let $w_g = (w_{1g}, ..., w_{Jg})'$. Then we can write (17) as a quadratic form:

$$S_{ML} = w_g' \tilde{A} w_g, \qquad \text{where} \qquad \tilde{A} = \begin{pmatrix} (1-\pi_1)\pi_1 & -\pi_1\pi_2 & \cdots & -\pi_1\pi_2 \\ -\pi_2\pi_1 & (1-\pi_2)\pi_2 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -\pi_J\pi_1 & -\pi_J\pi_2 & \cdots & (1-\pi_J)\pi_J \end{pmatrix}.$$

For the unbiased estimator we multiply by $\frac{N_I}{N_I - 1}$: $S_U = \frac{N}{N-1} S_{ML}$. We do not observe θ but its noisy estimate $\hat{\theta}$ and its sampling variance Σ. The unbiased estimate for the variance is therefore: $\widehat{\theta' B \theta} = \hat{\theta}' B \hat{\theta} - Tr(B\Sigma)$, where $Tr(\cdot)$ is the trace operator.

F Intermarriage and Location Effects

We evaluate the role of cultural assimilation in location effect heterogeneity by examining whether places that causally increase the probability of Russian-Israeli intermarriage also affect children's long-run economic outcomes. We begin by estimating Equation 2 on the sample of immigrants using an intermarriage indicator. In Table F.1, which reports the variance components of the intermarriage location effects, suggests that places do affect the likelihood of intermarriage, although these effects do not vary by parental income. Therefore, our final measure for intermarriage location effects is the slope parameter on years of exposure in each city—similar to Equation 2 but without the interaction with parental income (column 3). Lastly, in Table F.2, we present the coefficients from a weighted least squares regression of income rank location effects on the posterior mean intermarriage location effects.

Table F.1: Variance component for intermarriage location effects

	W/ income interaction		W/o income interaction
	Immigrants (1)	Natives (2)	Immigrants (3)
Std α	0.0037 (0.0003)	0.00004 (0.0056)	0.0041 (0.0003)
Std η	0.00004 (0.00001)	- -	

Note: This table reports the standard deviation of location effect on the probability for an intermarrige between Russians and immigrants. In column 1, the location effects vary linearly by parents' income rank, and in column 3, they don't. Column 1 gives the results for immigrants and column 2 for natives. Missing value indicates a negative variance component.

Table F.2: Intermarriage and child income ranks location effects

	Immigrants				Natives			
	θ_{25} (1)	θ_{25} (2)	θ_{75} (3)	θ_{75} (4)	θ_{25} (5)	θ_{25} (6)	θ_{75} (7)	θ_{75} (8)
Post. mean intermarrige effect	0.17 (0.18)	0.11 (0.18)	0.59 (0.20)	0.52 (0.20)	-0.07 (0.18)	-0.12 (0.18)	0.28 (0.17)	0.21 (0.16)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
# of cities	98	98	98	98	98	98	98	98

Note: This table reports the relationship between the location effects on income rank and the location effects on immigrants' intermarriage probability. In columns 1-2, the dependent variable is immigrants' location effects for p = 25, in columns 2-4, it's immigrants' location effects for p = 75, in columns 5-6, it's natives' location effects for p = 25, and in columns 7-8, it's natives' location effects for p = 75. Controls include population size in 2003, the diversity index, and locality welfare expenditure per capita.

G 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 $\hat{\theta}_g$ follows a normal distribution: $\hat{\theta}_g \sim \mathcal{N}(\theta_g, \Sigma_g)$, which can be justified by the central limit theorem with a growing number of families in each city. This section estimates the joint distribution of location effects while allowing the mean location effects to vary linearly with a few city-level covariates z_j. Ignoring the p subscript for simplicity, our model is described by:

$$\theta_{jg} = z_j' \beta_g + \nu_{jg}, \quad \nu_j | z_j, \Sigma \overset{iid}{\sim} G; \qquad \hat{\theta}_{jg} = \theta_{jg} + u_{jg}, \quad U_g | z, \Sigma \sim \mathcal{N}(0, \Sigma)$$

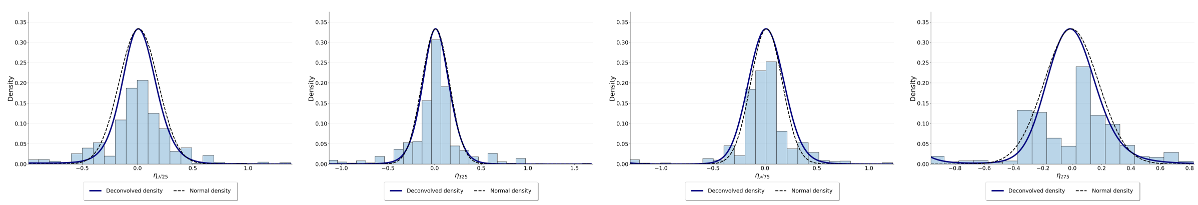
for $g \in \{\mathcal{N}, \mathcal{I}\}$, where $\nu_j = (\nu_{j\mathcal{N}}, \nu_{j\mathcal{I}})'$, $U_g = (u_{1g}, ..., u_{Jg})$, Σ is the $2J \times 2J$ sampling error covariance matrix with Σ_g on the diagonal and zeros in the off-diagonal, and z_j is a $p \times 1$ vector of city level covariates, with $z = (z_1, ..., z_p)$ the $J \times p$ design matrix. The prior distribution of immigrant-native location effects are defined by $\beta = (\beta_{\mathcal{N}}, \beta_{\mathcal{I}})'$ and G, the distribution of ν_j. We estimate β with city-size Weighted Least Squares regression and form each group's $J \times 1$ residual $r_{jg} = \hat{\theta}_{jg} - z_j' \beta_g$. Then, in the second step, we estimate the joint distribution of $r_j = (r_{j\mathcal{N}}, r_{j\mathcal{I}})'$.

Choice of z_j: Tables G.1 and G.2 report the estimates of β and corresponding variance components using diffrent covariates z_j. Our preferred model, presented in column (2), is chosen as the one that maximizes the explanatory power of location effects of both immigrants and natives.

G.1 Log-spline Estimator

We start by estimating the marginal distribution of each ν_{jg} nonparametrically using the deconvolution estimator from Efron (2016), where the prior distribution belongs to an exponential family, estimated flexibly by a fifth-order spline. Spline parameters are estimated via penalized maximum likelihood, weighted by the number of residents in each city. Following the approach taken in Kline et al. (2024), the penalization parameter is chosen to match the mean and method of moments variance estimate ν_j reported in panel (ii) in Tables G.1 and G.2. Figure G.1 plots the marginal deconvolved distribution of ν_{jg} separately for every $g \in \{\mathcal{N}, \mathcal{I}\}$ (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.

Figure G.1: Deconvolved density of the residuals of immigrants' and natives' location effects (η)



(a) ν_{25} Natives (b) ν_{25} immigrants (c) ν_{75} Natives (d) ν_{75} immigrants

Note: These figures plot the distribution of the residuals of immigrants' and natives' location effects. Residuals are the difference between the estimated location effects $\hat{\theta}$ and the $z_j'\beta$, where β is taken from column (2) in Tables G.1 and G.2. The solid blue line shows the weighted Efron (2016) deconvolved log-spline density. Histograms show the weighted estimated location effects. Dashed lines show the density of a normal distribution with the same mean and variance.

G.2 Normal Prior

Following the previous results, we assume that G follows a mean-zero normal distribution with variance Ω. Therefore, the joint distribution of $\hat{\theta}$ is given by: $\hat{\theta}|\beta, \Omega, \Sigma, z \sim \mathcal{N}(\mu(z), V)$, 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, and the posterior distribution we use for decision-making is:

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

G.3 Hierarchical Bayes

To assess the sensitivity of our EB policy estimates to hyperparameter noise due to finite sample bias, we add a robustness exercise based on a hierarchical Bayesian model. We assume the following flat priors: $\beta_{pg} \sim N(0, S)$, $\sigma_g \sim InverseGamma(S^{-1}, S^{-1})$, $\rho \sim$

Table G.1: Location effect hyperparameters ($p = 25$)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(i) Mean										
Intercept (I)	0.098	-0.111	-0.028	0.016	-0.084	-0.071	-0.108	-0.040	-0.081	-0.123
	(0.081)	(0.165)	(0.122)	(0.119)	(0.162)	(0.173)	(0.178)	(0.176)	(0.153)	(0.174)
Intercept (N)	0.389	0.331	0.366	0.362	0.328	0.413	0.357	0.387	0.386	0.357
	(0.072)	(0.147)	(0.105)	(0.103)	(0.147)	(0.245)	(0.247)	(0.246)	(0.168)	(0.181)
Diversity (I)		0.439			0.466		0.469			0.267
		(0.232)			(0.242)		(0.291)			(0.200)
Diversity (N)		0.122			0.118		0.122			0.122
		(0.241)			(0.240)		(0.242)			(0.243)
Welf. expnd. (I)	-0.003	-0.001	-0.002	0.001	0.001	0.001	-0.002	-0.002	0.001	0.001
	(0.012)	(0.012)	(0.012)	(0.015)	(0.015)	(0.013)	(0.013)	(0.013)	(0.012)	(0.013)
Welf. expnd. (N)	-0.048	-0.048	-0.048	-0.048	-0.048	-0.049	-0.049	-0.049	-0.049	-0.049
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.017)	(0.017)	(0.017)	(0.013)	(0.013)
log se (I)						-0.116	0.012	-0.016	-0.081	-0.072
						(0.084)	(0.112)	(0.108)	(0.051)	(0.050)
log se (N)						0.017	0.019	0.015	0.015	0.019
						(0.171)	(0.166)	(0.168)	(0.087)	(0.087)
Pop. size (I)	0.005	0.004	0.005			-0.001	0.005	0.004		
	(0.002)	(0.002)	(0.002)			(0.003)	(0.005)	(0.005)		
Pop. size (N)	-0.000	-0.001	-0.000			0.000	-0.000	0.000		
	(0.003)	(0.003)	(0.003)			(0.006)	(0.005)	(0.005)		
Share imm. (I)			0.600	0.606				0.548	0.336	
			(0.291)	(0.305)				(0.329)	(0.229)	
Share imm. (N)			0.111	0.110				0.109	0.109	
			(0.311)	(0.310)				(0.309)	(0.311)	
(ii) Rand. effects										
$\sigma_{\mathcal{N}}$	0.180	0.179	0.180	0.180	0.179	0.180	0.179	0.180	0.180	0.179
	(0.0471)	(0.0472)	(0.0470)	(0.0473)	(0.0472)	(0.0472)	(0.0473)	(0.0474)	(0.0476)	(0.0474)
$\sigma_{\mathcal{I}}$	0.166	0.154	0.154	0.162	0.160	0.158	0.154	0.154	0.156	0.155
	(0.0520)	(0.0517)	(0.0515)	(0.0516)	(0.0516)	(0.0520)	(0.0518)	(0.0514)	(0.0520)	(0.0521)
ρ	-0.012	-0.053	-0.038	-0.044	-0.060	0.005	-0.057	-0.032	-0.014	-0.025
	(0.670)	(0.920)	(1.100)	(0.682)	(0.876)	(1.170)	(0.764)	(0.806)	(0.744)	(0.746)
(iii) Implied total var.										
Natives	0.209	0.207	0.208	0.208	0.208	0.209	0.208	0.208	0.208	0.208
	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)
Immigrants	0.169	0.175	0.174	0.177	0.178	0.176	0.174	0.174	0.177	0.178
	(0.052)	(0.049)	(0.049)	(0.050)	(0.050)	(0.050)	(0.049)	(0.049)	(0.049)	(0.049)
Correlation	0.019	-0.003	0.004	-0.028	-0.033	0.001	-0.008	0.005	0.005	0.005
	(0.4819)	(0.3933)	(0.4256)	(0.4010)	(0.4954)	(0.7701)	(0.4225)	(0.4192)	(0.4060)	(0.3795)
	[-0.736, 0.832]	[-0.699, 0.713]	[-0.694, 0.728]	[-0.739, 0.665]	[-0.742, 0.672]	[-0.698, 0.700]	[-0.698, 0.707]	[-0.684, 0.714]	[-0.667, 0.701]	[-0.670, 0.710]
R^2 Natives	0.258	0.252	0.251	0.251	0.259	0.258	0.259	0.251	0.251	0.259
R^2 Immigrants	0.035	0.226	0.217	0.162	0.192	0.194	0.217	0.217	0.223	0.242
# of cities	98	98	98	98	98	98	98	98	98	98

Note: This table reports the estimates of the joint distribution of native-born and immigrant location effects for families in the 25th percentile of the income distribution. Panel (i) reports the mean native (N) and immigrant (I) location effects, a linear function of city-level covariates. Panel (ii) reports the standard deviation (σ) and the correlation (ρ) of the ν_j, the random effects, and panel (iii) reports the implied total standard deviation, 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 the group specific ratio between variance share $\frac{\beta_g^2 \mathbb{V}(z)}{\beta_g^2 \mathbb{V}(z) + \sigma_g^2}$. Panel (i) was estimated by a city-size weighted least squares regression. In panel (i), robust standard errors are reported in parentheses. In panels (ii)-(iii), parentheses report the parametric bootstrapped standard errors, and square brackets report parametric bootstrapped equal-tailed confidence intervals.

$Uniform[-1, 1]$, experimenting with two values for $S \in \{5, 1000\}$. We estimate the model by sampling from the posterior distribution of location effects and hyperparameters, conditional on the data, using Hamiltonian Monte Carlo (HMC), a modern Markov chain Monte Carlo (MCMC) method implemented in the Python NumPyro package.

As can be seen in Table G.3, which reports the EB and hierarchical Bayes (HB) estimates of the mean and variance for immigrants and natives, the two methods result in very similar estimates, especially for the less informative prior when $S = 1000$. In line with that, Table G.4 reports the list of selected cities under the hierarchical Bayes model with the uninformative prior ($S = 1000$) and shows that the list of selected cities aligns closely with the list selected via the EB procedure we take in the paper.

Table G.2: Location effect hyperparameters ($p = 75$)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(i) Mean										
Intercept (I)	0.291	-0.002	0.106	0.149	0.023	-0.104	-0.104	-0.070	-0.014	-0.063
	(0.093)	(0.179)	(0.131)	(0.140)	(0.181)	(0.196)	(0.197)	(0.208)	(0.157)	(0.183)
Intercept (N)	0.123	-0.054	0.047	0.069	-0.042	-0.090	-0.397	-0.291	-0.162	-0.247
	(0.067)	(0.107)	(0.084)	(0.088)	(0.110)	(0.228)	(0.271)	(0.275)	(0.173)	(0.176)
Diversity (I)		0.617			0.643		0.119			0.344
		(0.289)			(0.299)		(0.515)			(0.287)
Diversity (N)		0.371			0.383		0.436			0.397
		(0.190)			(0.188)		(0.203)			(0.189)
Welf. expnd. (I)	-0.015	-0.013	-0.014	-0.011	-0.010	-0.001	-0.003	-0.004	-0.008	-0.008
	(0.016)	(0.016)	(0.016)	(0.020)	(0.019)	(0.018)	(0.020)	(0.020)	(0.015)	(0.015)
Welf. expnd. (N)	-0.033	-0.031	-0.032	-0.031	-0.030	-0.023	-0.017	-0.018	-0.024	-0.024
	(0.012)	(0.013)	(0.012)	(0.014)	(0.014)	(0.014)	(0.015)	(0.015)	(0.011)	(0.010)
log se (I)						-0.247	-0.211	-0.185	-0.118	-0.113
						(0.097)	(0.178)	(0.150)	(0.060)	(0.062)
log se (N)						-0.141	-0.207	-0.208	-0.122	-0.111
						(0.152)	(0.145)	(0.159)	(0.084)	(0.079)
Pop. size (I)	0.005	0.004	0.005			-0.007	-0.006	-0.004		
	(0.003)	(0.003)	(0.003)			(0.005)	(0.009)	(0.008)		
Pop. size (N)	0.002	0.002	0.002			-0.002	-0.004	-0.004		
	(0.002)	(0.002)	(0.002)			(0.004)	(0.004)	(0.004)		
Share imm. (I)			0.882	0.887				0.307	0.515	
			(0.378)	(0.395)				(0.588)	(0.348)	
Share imm. (N)			0.357	0.360				0.476	0.425	
			(0.262)	(0.265)				(0.291)	(0.271)	
(ii) Rand. effects										
$\sigma_{\mathcal{N}}$	0.161	0.152	0.157	0.159	0.153	0.159	0.146	0.152	0.153	0.148
	(0.0449)	(0.0451)	(0.0448)	(0.0454)	(0.0449)	(0.0449)	(0.0451)	(0.0451)	(0.0448)	(0.0446)
$\sigma_{\mathcal{I}}$	0.211	0.191	0.190	0.196	0.195	0.186	0.186	0.185	0.186	0.187
	(0.0468)	(0.0473)	(0.0478)	(0.0472)	(0.0474)	(0.0483)	(0.0479)	(0.0479)	(0.0476)	(0.0478)
ρ	0.420	0.327	0.366	0.387	0.343	0.340	0.341	0.326	0.340	0.331
	(0.683)	(0.675)	(0.623)	(0.595)	(0.777)	(0.584)	(1.622)	(1.232)	(0.971)	(0.790)
(iii) Implied total var.										
Natives	0.178	0.177	0.178	0.176	0.177	0.179	0.181	0.180	0.178	0.178
	(0.042)	(0.040)	(0.041)	(0.042)	(0.041)	(0.041)	(0.039)	(0.040)	(0.040)	(0.039)
Immigrants	0.216	0.222	0.222	0.223	0.224	0.232	0.230	0.229	0.225	0.225
	(0.046)	(0.040)	(0.041)	(0.041)	(0.041)	(0.038)	(0.038)	(0.038)	(0.039)	(0.039)
Correlation	0.453	0.459	0.453	0.446	0.454	0.338	0.511	0.456	0.459	0.490
	(0.3608)	(0.3577)	(0.3695)	(0.7962)	(0.3596)	(0.2771)	(0.2962)	(0.3051)	(0.3459)	(0.3580)
	[-0.071, 1.222]	[-0.004, 1.179]	[-0.020, 1.170]	[-0.039, 1.162]	[-0.010, 1.176]	[-0.113, 0.915]	[0.085, 1.171]	[0.020, 1.105]	[0.013, 1.151]	[0.049, 1.195]
R^2 Natives	0.182	0.263	0.222	0.184	0.253	0.211	0.349	0.287	0.261	0.309
R^2 Immigrants	0.046	0.260	0.268	0.227	0.242	0.357	0.346	0.347	0.317	0.309
# of cities	98	98	98	98	98	98	98	98	98	98

Note: This table reports the estimates of the joint distribution of native-born and immigrant location effects. Panel (i) reports the mean native (N) and immigrant (I) location effects, a linear function of city-level covariates. Panel (ii) reports the standard deviation (σ) and the correlation (ρ) of the ν_j, the random effects, and panel (iii) reports the implied total standard deviation, 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 the group specific ratio between variance share $\frac{\beta_g^2 \mathbb{V}(z)}{\beta_g^2 \mathbb{V}(z) + \sigma_g^2}$. Panel (i) was estimated by a city-size weighted least squares regression. In panel (i), robust standard errors are reported in parentheses. In panels (ii)-(iii), parentheses report the parametric bootstrapped standard errors, and square brackets report parametric bootstrapped equal-tailed confidence intervals.

H Full List of the Posterior Mean Location Effects

Table H.1: Forecast of location effects

		$p = 25$		$p = 75$	
Loc. Code	Name	Posterior mean imm. (1)	Posterior mean natives (2)	Posterior mean imm. (3)	Posterior mean natives (4)
246	Netivot	0.265	-0.177	0.244	-0.005
7100	Ashqelon	0.255	-0.020	0.311	0.108
70	Ashdod	0.234	0.069	0.467	0.112
6200	Bat Yam	0.183	0.233	0.461	0.397
2560	Arad	0.162	0.064	0.142	0.124
4100	Qazrin	0.157	-0.005	0.148	0.060
9100	Nahariyya	0.150	-0.185	0.072	-0.225
1139	Karmi'el	0.143	0.142	0.075	0.104

Continued on next page

Table H.1: Forecast of location effects (cont.)

		$p = 25$		$p = 75$	
		Posterior mean imm.	Posterior mean natives	Posterior mean imm.	Posterior mean natives
Loc. Code	Name	(1)	(2)	(3)	(4)
2200	Dimona	0.140	0.077	0.407	0.190
2100	Tirat Karmel	0.130	-0.124	0.087	-0.042
4000	Haifa	0.130	0.149	-0.054	0.112
3640	Qarne Shomeron	0.105	0.089	0.192	0.110
3570	Ari'el	0.104	-0.029	0.079	0.018
7600	Akko	0.103	-0.231	0.313	0.077
240	Yoqne'am Illit	0.101	0.067	0.137	0.131
1137	Qiryat Ye'arim	0.100	-0.040	0.100	0.019
8300	Rishon Leziyyon	0.100	0.167	0.280	0.174
99	Mizpe Ramon	0.095	-0.028	0.103	0.014
8400	Rehovot	0.094	-0.078	-0.037	-0.055
1034	Qiryat Mal'akhi	0.090	-0.250	0.140	-0.120
6600	Holon	0.090	0.148	0.167	0.157
5000	Tel-Aviv	0.088	-0.193	0.105	-0.033
31	Ofaqim	0.087	-0.158	-0.034	0.050
2630	Qiryat Gat	0.085	0.171	0.244	0.202
3780	Betar Illit	0.079	0.217	0.069	0.059
874	Migdal Haemeq	0.075	-0.232	-0.067	-0.151
7700	Afula	0.070	0.003	0.209	0.144
76*	Gush Etzion	0.069	0.095	0.111	0.097
73*	Mateh Binyamin	0.067	0.057	0.076	0.072
2660	Yavne	0.063	0.020	0.094	-0.012
8500	Ramla	0.058	-0.043	0.218	0.064
1063	Ma'alot-tarshiha	0.056	0.100	0.093	0.115
7000	Lod	0.044	-0.046	0.119	0.020
1066	Bene Ayish	0.044	0.096	0.072	0.103
1*	Upper Galilee	0.040	-0.163	-0.048	-0.140
2500	Nesher	0.038	-0.022	-0.004	-0.039
1020	Or Aqiva	0.034	-0.158	0.135	-0.013
9500	Qiryat Bialik	0.033	-0.036	-0.038	0.018
8700	Ra'annana	0.033	0.001	0.080	0.031
2800	Qiryat Shemona	0.032	-0.058	-0.019	-0.015
38*	Eshkol	0.030	0.033	-0.055	-0.029
6700	Tiberias	0.027	0.068	0.142	0.098
2400	Or Yehuda	0.025	-0.079	0.271	0.075
812	Shelomi	0.024	-0.380	-0.052	-0.211
9000	Be'er Sheva	0.021	0.037	-0.058	-0.033
3650	Efrata	0.013	0.195	0.084	0.207
6300	Giv'atayim	0.008	-0.074	0.086	-0.026
6100	Bene Beraq	0.008	-0.305	0.120	-0.143
9600	Qiryat Yam	0.006	0.083	0.232	0.141
7900	Petah Tiqwa	0.006	-0.020	0.106	0.024
565	Azor	0.005	-0.014	-0.034	-0.008
1061	Nazerat Illit	0.002	-0.084	0.005	-0.031
2034	Hazor Hagelilit	0.001	0.032	-0.007	0.016
7400	Netanya	0.000	-0.001	0.191	0.084
1031	Sederot	-0.006	-0.083	0.160	0.044
6900	Kefar Sava	-0.013	0.060	-0.094	0.048
8200	Qiryat Motzkin	-0.017	0.177	0.035	0.179
831	Yeroham	-0.027	0.022	-0.074	-0.021
72*	Shomron	-0.032	0.193	-0.016	0.126
2530	Be'er Ya'aqov	-0.038	-0.012	-0.114	-0.020
1224	Kokhav Ya'ir	-0.039	0.195	0.005	0.071
9400	Yehud	-0.040	-0.010	-0.102	-0.079
168	Kefar Yona	-0.042	-0.087	-0.135	-0.194
3616	Ma'ale Adummim	-0.056	0.140	-0.081	0.039
7200	Nes Ziyyona	-0.058	-0.002	-0.109	-0.101
1015	Mevasseret Ziyyon	-0.060	0.229	-0.071	-0.019
2620	Qiryat Ono	-0.062	-0.027	-0.107	-0.142
469	Qiryat Eqron	-0.062	0.162	-0.173	-0.015
9700	Hod Hasharon	-0.064	0.177	-0.047	0.049
15*	Hof HaCarmel	-0.072	-0.210	-0.150	-0.182
26*	Mateh Yehuda	-0.077	0.074	-0.117	-0.094
6*	Emek HaYarden	-0.082	-0.190	-0.221	-0.218

Continued on next page

Table H.1: Forecast of location effects (cont.)

		$p = 25$		$p = 75$	
		Posterior mean imm.	Posterior mean natives	Posterior mean imm.	Posterior mean natives
Loc. Code	Name	(1)	(2)	(3)	(4)
9200	Bet She'an	-0.086	-0.007	-0.180	-0.135
8000	Zefat	-0.089	0.062	-0.038	0.023
3611	Qiryat Arba	-0.091	-0.261	-0.205	-0.245
8600	Ramat Gan	-0.100	0.151	-0.196	0.028
6800	Qiryat Atta	-0.104	0.069	-0.095	0.006
2600	Elat	-0.105	0.125	0.114	0.159
166	Gan Yavne	-0.106	0.018	-0.125	0.032
71*	Golan	-0.107	0.065	-0.175	-0.065
56*	Misgav	-0.110	0.067	-0.141	-0.057
9300	Zikhron Ya'aqov	-0.115	0.015	-0.130	0.022
6500	Hadera	-0.119	0.247	-0.006	0.089
2650	Ramat Hasharon	-0.125	0.100	-0.121	-0.007
7*	Emek HaMaayanot	-0.130	-0.002	-0.193	-0.115
9*	Jezreel Valley	-0.134	0.087	-0.187	-0.014
2640	Rosh Haayin	-0.137	0.150	-0.139	0.004
6400	Herzliyya	-0.153	-0.164	-0.223	-0.102
681	Giv'at Shemu'el	-0.155	-0.085	-0.252	-0.116
33*	Be'er Tuvia	-0.156	-0.033	-0.202	-0.104
2610	Bet Shemesh	-0.159	-0.012	-0.035	0.207
30*	Gezer	-0.170	0.095	-0.175	0.024
42*	Merhavim	-0.170	-0.351	-0.407	-0.447
8*	Gilboa	-0.173	0.160	-0.156	-0.011
16*	Hefer Valley	-0.180	-0.064	-0.243	-0.133
2550	Gedera	-0.194	-0.005	-0.239	-0.099
4*	Mateh Asher	-0.207	0.001	-0.359	-0.203
7800	Pardes Hanna-karkur	-0.216	-0.115	-0.297	-0.058

Note: This table presents the posterior mean location effects for immigrant and native born Israeli children. Columns 1-2 report the posterior mean location effects for children with parents at the 25th percentile of the national income distribution, and columns 3-4 report the effects for children with parents at the 75th percentile of the income distribution. Locations are listed with their name and location code, where an asterisk marks regional council codes. The table is sorted by the posterior mean of immigrants with $p = 25$.

I Unified Policies

Many public policies, including neighborhood recommendation systems, face horizontal equity constraints, which require treating people equally irrespective of protected attributes such as race or ethnicity. Below are examples from different regions that highlight why such unified (non-personalized) policies are widespread. Horizontal equity, meaning equal treatment across protected groups, is not just normative but frequently legally mandated across diverse jurisdictions. A few non-exhaustive examples include:

- US Fair Housing Act, 1968:[36] Federal law prohibits housing discrimination based on race, color, religion, sex, family status, disability, or national origin. Among other provisions, it prohibits housing agencies from discriminating against or restricting access to public housing based on these protected characteristics. This prohibition applies to private, publicly funded, and public housing alike.

  A landmark case, Davis et al. v. The St. Louis Housing Authority (1952), illustrates

[36] See https://www.hud.gov/helping-americans/fair-housing-act-overview?utm_source=chatgpt.com.

Table G.3: Hyperparameters estimates, EB and full Bayes

	$\mu(z)$ Immigrants			$\mu(z)$ Natives			Variance			
	EB	HB (S=5)	HB (S=1000)	EB	HB (S=5)	HB (S=1000)		EB	HB (S=5)	HB (S=1000)
	(1)	(2)	(3)	(4)	(5)	(6)		(7)	(8)	(9)
Intercept	-0.228	-0.173	-0.176	0.197	0.191	0.137	σ_N	0.179	0.197	0.122
	(0.165)	(0.193)	(0.156)	(0.147)	(0.198)	(0.165)		(0.047)	(0.029)	(0.027)
Population	0.004	0.004	0.005	-0.001	-0.001	-0.001	σ_I	0.154	0.231	0.143
	(0.002)	(0.005)	(0.003)	(0.003)	(0.005)	(0.004)		(0.052)	(0.037)	(0.042)
Diversity	0.439	0.408	0.402	0.122	0.122	0.274	ρ	-0.053	-0.256	-0.159
	(0.232)	(0.294)	(0.229)	(0.241)	(0.307)	(0.258)		(0.920)	(0.357)	(0.328)
Welf. expnd.	-0.001	-0.010	-0.010	-0.048	-0.045	-0.049		[-0.699, 0.713]	[-0.816, 0.506]	[-0.737, 0.507]
	(0.012)	(0.022)	(0.017)	(0.014)	(0.027)	(0.022)				

Note: This table reports the estimates of the hyperparameters governing the distribution of location effects for immigrants and natives. Columns 1–3 and 4–6 report the coefficients for the mean location effects of immigrants and natives, respectively, where the mean is modeled as a linear function of city-level covariates: population size (in thousands), diversity index, and municipal welfare expenditure per capita. Columns 7–9 report the parameters for the variance of the random effects. The EB hyperparameters were estimated via weighted least squares (for the mean) and the method of moments (for the variance), while the full hierarchical Bayes (HB) hyperparameters were estimated via Hamiltonian Monte Carlo (HMC). Values in parentheses are standard errors for EB estimates and posterior standard deviations for FB estimates. Square brackets report parametric bootstrapped equal-tailed confidence intervals for the EB estimate of ρ, and the posterior credible interval for the FB estimate of ρ.

Table G.4: Selected cities based on minimax criterion, $K = 10$, full Bayes estimates

	Posterior mean					
	Native-born	Imm.	Average	$E[\max\{\vartheta_{\mathcal{N}}, \vartheta_{\mathcal{I}}\}\|\mathcal{Y}]$	Worse than status-quo	Selected by minimax ($\mathcal{I}/\mathcal{N}$/city)
Loc. name	(1)	(2)	(3)	(4)	(5)	(6)
Bat Yam	128.2	145.3	139.9	198.5	0.004	Yes (631.7)
Rishon Leziyyon	245.3	211.4	217.7	280.1	0.008	Yes (631.7)
Holon	270.6	264.8	265.8	325.0	0.043	Yes (631.7)
Haifa	219.1	331.4	300.9	344.6	0.115	Yes (631.7)
Qiryat Gat	286.0	213.7	237.3	368.5	0.240	Yes (631.7)
Ashdod	46.3	381.7	258.9	381.9	0.234	Yes (631.7)
Karmi'el	237.3	316.3	285.3	396.2	0.301	Yes (631.7)
Bene Ayish	321.7	282.4	305.2	449.1	0.433	Yes (631.7)
Be'er Sheva	401.4	400.8	401.0	456.1	0.393	Yes (631.7)
Kokhav Ya'ir	405.8	162.0	207.3	472.4	0.490	
Ma'alot-tarshiha	342.5	383.2	363.7	497.8	0.533	
Qiryat Motzkin	437.3	311.4	339.1	500.6	0.571	Yes (631.7)
Arad	283.7	436.3	372.3	503.6	0.545	
Yoqne'am Illit	356.8	403.7	391.3	517.7	0.587	
Ari'el	317.7	428.3	380.7	518.1	0.569	

Note: This table reports a list of 15 Israeli cities sorted by within-city posterior immigrant-native maximum regret. Posterior estimates are based on a full hierarchical Bayesian model reported in columns 3, 6, and 9 in Table G.3. Columns 1-3 report the posterior mean regret of native-born children, immigrants, and the average. Column 4 reports the posterior maximum $(\mathcal{N}, \mathcal{I})$. Column 5 reports the posterior probability that the location effects of immigrants or of natives are lower than their average effect under the status quo sorting patterns. Column 6 reports the cities selected based on the minimax $(\mathcal{N}/\mathcal{I}/city)$ policy that ranks lists of 10 cities based on their posterior maximum $(\mathcal{N}, \mathcal{I})/city$ regret (reported in parentheses).

the Act's power: the court ruled that the Authority unlawfully limited certain public housing units to white residents only and enjoined them from denying units to qualified African American applicants based solely on race.

- European Union (EU AI Act): The EU Artificial Intelligence Act (Regulation (EU) 2024/1689, hereafter EU AI Act) prohibits AI systems from classifying or targeting individuals based on sensitive personal traits such as race, religion, or sexual orientation. Systems that manipulate decisions or exploit such attributes are banned to ensure non-discriminatory outcomes.

- Israel's legal framework related to housing and anti-discrimination: The Prohibition of Discrimination in Products, Services and Entry into Places of Entertainment and

Public Places Law (2000) forbids discrimination by providers of goods, services, or access to public venues on the basis of race, religion, nationality, and other attributes. Although not specific to housing, this law underscores the broader principle of non-discrimination in public contexts.

In discussions with Israeli government officials who expressed interest in the CMTO experiment, they noted that personalized policies would likely be politically infeasible. After reviewing evidence of high heterogeneity in location effects in Israel, they expressed concern that a fully personalized policy would conflict with political and social norms favoring uniform treatment.

J Policy Estimation and Simulation

Estimation of ϑ and its prior distribution: The decision maker evaluates the contribution of each location j for each group $g \in \{\mathcal{N}, \mathcal{I}\}$ by its regret, i.e., relative to the expected value under the first-best policy that allows discrimination:

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

where $\theta^*(\delta^*_{gK}, K) \equiv \frac{1}{K}\sum_{j=1}^{J} \mathbb{E}[\theta_{jg}\delta^*_{jgK}]$ is group g's expected long-run effect of selected cities under the first-best.

While we directly estimate θ_{jg}, ϑ_{jgK} depends both on the distribution of θ_{jg} and on $\theta^*(\delta^*_{gK}, K)$ which is a function of the distribution of the K-th ordered statistic. To estimate ϑ_{jgK}, we proceed as follows: using the distribution of θ (estimated in Appendix Section G), for every K we approximate $\theta^*(\delta^*_{gK}, K)$ via Monte Carlo simulation with 10,000 draws from the prior distribution. Appendix Table J.1 reports the estimates of $\theta^*(\delta^*_{gK}, K)$, and their standard errors for K=1,..,10.

Finally, we set $\hat{\vartheta}_{jgK} = \theta^*(\delta^*_{gK}, K) - \hat{\theta}_{jg}$, where for every K, we plug the corresponding $\theta^*(\delta^*_{gK}, K)$ from Table J.1, treated as contant measured with no noise. Then, given $\hat{\vartheta}_{jgK}$, and their variance covariance matrix Σ, we estimate the prior distribution of ϑ_{jgK} following the same steps depicted in Appendix Section G.

Estimation of minimax policies: The model in Section 10.3 describes decision rules that require estimating the posterior expectation of the maximum. We estimate these policies using a simulation: we draw 1000 samples from the posterior distribution of $\vartheta|\hat{\vartheta}, \Sigma, z$, and within each draw, we estimate the minimax policies. For the $(\mathcal{N}, \mathcal{I}, city)$ policy, we approximate the optimal decision rule using a beam search, which restricts the search space to the top-30 most promising candidate city combinations at each step,

Table J.1: Estimates for $\theta^*(\delta^*_{gK}, K)$, p=25

	Native-born (1)	Immigrants (2)
K=1	0.5217	0.6181
	(0.0007)	(0.0008)
K=2	0.4911	0.5823
	(0.0006)	(0.0007)
K=3	0.4693	0.5570
	(0.0005)	(0.0006)
K=4	0.4523	0.5372
	(0.0004)	(0.0005)
K=5	0.4381	0.5208
	(0.0004)	(0.0005)
K=6	0.4260	0.5066
	(0.0004)	(0.0004)
K=7	0.4152	0.4942
	(0.0004)	(0.0004)
K=8	0.4056	0.4830
	(0.0003)	(0.0004)
K=9	0.3968	0.4729
	(0.0003)	(0.0004)
K=10	0.3887	0.4635
	(0.0003)	(0.0004)

Note: This table reports the estimates of $\theta^*(\delta^*_{gK}, K) \equiv \frac{1}{K}\sum_{j=1}^{J} \mathbb{E}[\theta_{jg}\delta^*_{jgK}]$, which represents each group $g \in \{\mathcal{N}, \mathcal{I}\}$'s expected long-run effect of the selected cities under the first-best personalized policy, for $K = 1, \ldots, 10$. Estimates are obtained via Monte Carlo simulation with 10,000 draws. Standard errors, reported in parentheses, are calculated as the Monte Carlo sample standard deviation divided by $\sqrt{10{,}000}$.

based on interim regret values. This allows us to efficiently approximate the minimax regret without exhaustively evaluating all the $\binom{98}{10}$ possible subsets.

Simulation of policy counterfactuals: To produce Figure 6, we draw ϑ_j for the 98 cities and groups from its prior distribution (see above for an explanation of how we estimate this distribution). Given a fixed set of cities, we then simulate 2000 iid draws of the noise from a normal distribution with mean zero and variance Σ and compute $\hat{\vartheta}_j$. Then, to compute the EB policies, we draw 1000 samples from the posterior distribution of each *simulation* and follow the procedure mentioned above. After computing the policy for every simulation, we calculate each group's average and max regret among selected cities and collect the mean value, averaging across the 2000 simulations. To get the performance of an average DGP, we repeat this entire procedure 100 times and report the average.

K Bayes Minimax Policies

This section provides a detailed explanation and proofs for the Empirical-Bayes-minimax policies introduced in Section 10.3. Our models are comprised of a decision maker whose goal is to select K cities based on their long-run effects on children's outcomes. The decision maker observes the estimates for the normalized location effects (regret) $\hat{\vartheta} = (\hat{\vartheta}_{1,\mathcal{N}}, ..., \hat{\vartheta}_{J,\mathcal{N}}, \hat{\vartheta}_{1,\mathcal{I}}, ..., \hat{\vartheta}_{J,\mathcal{I}})'$ and their samplaing variance Σ, which we collect into the array $\mathcal{Y} = (\hat{\vartheta}, \Sigma)$. After observing $\mathcal{Y}$, the decision maker selects K cities, an action we call $\delta \in \{0,1\}^J$ such that $\sum_j \delta_j = K$, where the decision rule $\boldsymbol{\delta}(.) : \mathbb{P}(\mathcal{Y}) \to \{0,1\}^J$ maps realizations of the data into recommendations.

The decision maker's loss $\mathcal{L}(\vartheta, \delta, \pi(\delta))$ depends on ϑ, the choice of δ, and inportantly, $\pi(\delta) = (\pi_{1\mathcal{N}}(\delta), ...\pi_{J\mathcal{N}}(\delta), \pi_{1\mathcal{I}}(\delta), ..., \pi_{J\mathcal{I}}(\delta))'$, which describes how families sort into cities when facing the policy δ. Throughout, we assume that the decision maker is ambiguous about $\pi(\delta)$, i.e., holds no prior information on the distribution of $\pi(\delta)$. While $\mathcal{Y}$ can be used to learn about ϑ, it contains no information about the true value of $\pi(\delta)$.

K.1 Who Shows Up?

The first model considers a decision maker who is interested in choosing K cities and building a single housing unit in each city. The decision maker's loss function is therefore given by:

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

where $\pi_{j\mathcal{I}}(\delta) \in [0, 1]$ decribes the probability that an immigrant family moves to city j given the list of selected cities δ.

Known ϑ: If the decisionmaker knows ϑ, then the data are irrelevant and the decision rule is directly a function of ϑ. To handle the ambiguity about $\pi(\delta)$, the decision maker evaluates actions by their maximum loss:

$$\mathcal{L}^{(\mathcal{N},\mathcal{I})}(\vartheta, \delta) = \max_{\pi(\delta)} \mathcal{L}(\vartheta, \delta, \pi(\delta)) = \sum_j \delta_j \max_{\pi_{j\mathcal{I}}(\delta)\in\{0,1\}} \left\{\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}K} + (1 - \pi_{j\mathcal{I}}(\delta))\vartheta_{j\mathcal{N}K}\right\},$$

and the optimal policy satisfies:

$$\boldsymbol{\delta}(\vartheta) = \underset{\delta,\ \text{s.t} \sum \delta_j = K}{\arg\min}\ \mathcal{L}^{(\mathcal{N},\mathcal{I})}(\vartheta, \delta),$$

which implies that for every city j, the optimal policy given ϑ is:

$$\delta^*_{jK} = \mathbb{1}\{\max\{\vartheta_{j\mathcal{I}K}, \vartheta_{j\mathcal{N}K}\} \leq \kappa_K\},$$

where κ_K is the maximum value such that there are exactly K cities with $\max\{\vartheta_{j\mathcal{I}K}, \vartheta_{j\mathcal{N}K}\} \leq \kappa_K$.

Proof. First note that for every j, $\max_{\pi_{j\mathcal{I}}(\delta)\in\{0,1\}} \{\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}K} + (1 - \pi_{j\mathcal{I}}(\delta))\vartheta_{j\mathcal{N}K}\} = \max\{\vartheta_{j\mathcal{I}K}, \vartheta_{j\mathcal{N}K}\}$. Let $\delta \neq \delta^*_{jK}$ with $\sum_j \delta_j = K$. Then for every city j with $\delta_j = 1 \neq \delta^*_{jK}$ we have $\max\{\vartheta_{j\mathcal{I}K}, \vartheta_{j\mathcal{N}K}\} > \kappa_K$, and for every city l with $\delta_l = 0 \neq \delta^*_{lK}$ we have $\max\{\vartheta_{l\mathcal{I}K}, \vartheta_{l\mathcal{N}K}\} \leq \kappa_K$, and therefore $\mathcal{L}^{(\mathcal{N},\mathcal{I})}(\vartheta, \delta) > \mathcal{L}^{(\mathcal{N},\mathcal{I})}(\vartheta, \delta^*_K)$. $\square$

Unknown ϑ: When ϑ is unknown, the decision maker uses the data $\mathcal{Y}$ to learn about the true ϑ before choosing a policy δ. Therefore, the decision maker's risk (expected loss) is defined by:

$$\mathcal{R}^{(\mathcal{N},\mathcal{I})}(\delta) = \mathbb{E}\left[\mathcal{L}^{(\mathcal{N},\mathcal{I})}(\vartheta,\delta)\middle|\mathcal{Y}\right] = \sum_j \delta_j \mathbb{E}\left[\max_{\pi_j}\left\{\pi_{j\mathcal{I}}\vartheta_{j\mathcal{I}K} + (1-\pi_{j\mathcal{I}})\vartheta_{j\mathcal{N}K}\right\}\middle|\mathcal{Y}\right], \quad (18)$$

where $\mathbb{E}[\cdot|\mathcal{Y}]$ denotes expectation with respect to the posterior distribution of ϑ given $\mathcal{Y}$.[37] Optimal policy satisfies:

$$\boldsymbol{\delta}(\mathcal{Y}) = \underset{\delta,\ \sum\delta_j = K}{\arg\min}\ \mathcal{R}^{(\mathcal{N},\mathcal{I})}(\delta)$$

Which implies that for every city j, the optimal policy given the ebvidence $\mathcal{Y}$ is:

$$\delta_{jK}^{*(\mathcal{N},\mathcal{I})} = \mathbb{1}\{\mathbb{E}\left[\max\{\vartheta_{j\mathcal{I}K},\vartheta_{j\mathcal{N}K}\}|\mathcal{Y}\right] \le \kappa_K\},$$

Proof. Since $\max_{\pi_{j\mathcal{I}}(\delta)\in\{0,1\}}\left\{\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}K} + (1-\pi_{j\mathcal{I}}(\delta))\vartheta_{j\mathcal{N}K}\right\} = \max\{\vartheta_{j\mathcal{I}K},\vartheta_{j\mathcal{N}K}\}$, then $\mathbb{E}[\max_{\pi_j}\left\{\pi_{j\mathcal{I}}\vartheta_{j\mathcal{I}K} + (1-\pi_{j\mathcal{I}})\vartheta_{j\mathcal{N}K}\right\}|\mathcal{Y}] = \mathbb{E}\left[\max\{\vartheta_{j\mathcal{I}K},\vartheta_{j\mathcal{N}K}\}|\mathcal{Y}\right]$. Then, similar to the argument in K.1, let $\delta \ne \delta_K^{*(\mathcal{N},\mathcal{I})}$ with $\sum_j \delta_j = K$. For every city j with $\delta_j = 1 \ne \delta_{jK}^{*(\mathcal{N},\mathcal{I})}$, we have $\mathbb{E}\left[\max\{\vartheta_{j\mathcal{I}K},\vartheta_{j\mathcal{N}K}\}|\mathcal{Y}\right] > \kappa_K$, and for every city l with $\delta_l = 0 \ne \delta_{lK}^{*(\mathcal{N},\mathcal{I})}$, we have $\mathbb{E}\left[\max\{\vartheta_{l\mathcal{I}K},\vartheta_{l\mathcal{N}K}\}|\mathcal{Y}\right] \le \kappa_K$. Therefore $\mathcal{R}^{(\mathcal{N},\mathcal{I})}(\delta) > \mathcal{R}^{(\mathcal{N},\mathcal{I})}(\delta_{jK}^{*(\mathcal{N},\mathcal{I})})$. $\square$

K.2 Who Shows Up and Where Do They Go?

The second model considers a decision maker who would like to provide housing voucher recipients with a list of K recommended cities. In this model, the decision maker's loss function is given by:

$$\mathcal{L}(\vartheta,\delta,\pi(\delta)) = \sum_{j=1}^{J}\left[\delta_j(\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}K} + \pi_{j\mathcal{N}}(\delta)\vartheta_{j\mathcal{N}K})\right], \quad (19)$$

where (with a slight abuse of notation), $\pi_{jg}(\delta) \in [0,1]$ is the share of group $g \in \{\mathcal{N},\mathcal{I}\}$ families sorting into city j when city j is selected, such that $\sum_j(\pi_{j\mathcal{N}}(\delta) + \pi_{j\mathcal{I}}(\delta)) = 1$.

Known ϑ: Similarly, if ϑ is known, then the data are irrelevant and the decision rule is directly a function of ϑ. To handle the ambiguity about $\pi(\delta)$, the decision maker evaluates actions by their maximum loss:

$$\mathcal{L}^{(\mathcal{N},\mathcal{I},city)}(\vartheta,\delta) = \max_{\pi(\delta)}\mathcal{L}(\vartheta,\delta,\pi(\delta)) = \max_{\substack{\pi_{j\mathcal{I}}(\delta),\pi_{j\mathcal{N}}(\delta)\in\{0,1\},\text{ s.t}\\ \sum_j\left(\pi_{j\mathcal{N}}(\delta)+\pi_{j\mathcal{I}}(\delta)\right)=1}}\left\{\sum_j \delta_j(\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}K} + \pi_{j\mathcal{N}}(\delta))\vartheta_{j\mathcal{N}K})\right\},$$

[37]Formmaly, risk is defined as $\int\int \mathcal{L}^{(\mathcal{N},\mathcal{I})}(\vartheta,\boldsymbol{\delta}(x))f(x;\vartheta)dxdG(\vartheta)$. Standard arguments (e.g., in Wald (1950)) imply that minimizing it is equivalent to minimizing the posterior maximum risk in Equation (22).

and the optimal policy satisfies:

$$\boldsymbol{\delta}(\vartheta) = \underset{\delta,\ \text{s.t} \sum \delta_j = K}{\arg\min} \mathcal{L}^{(\mathcal{N},\mathcal{I},city)}(\vartheta, \delta).$$

Therefore, the optimal policy is:

$$\delta^*_{jK} = \underset{\delta \text{ s.t } \sum_j \delta_j = K}{\arg\min} \left\{ \max(\{\vartheta_{j\mathcal{I}K}, \vartheta_{j\mathcal{N}K}\}_{j\in S(\delta)}) \right\}, \tag{20}$$

where $S(\delta) = \{j : \delta_j = 1\}$ is the set of recommended cities.

Proof. It's immediate to see that when $\sum_{j=1}^{J} (\pi_{j\mathcal{N}}(\delta) + \pi_{j\mathcal{I}}(\delta)) = 1$, then

$$\max_{\pi_{j\mathcal{I}}(\delta)\in\{0,1\}} \left\{ \sum_j \delta_j (\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}K} + \pi_{j\mathcal{N}}(\delta))\vartheta_{j\mathcal{N}K}) \right\} = \max(\{\vartheta_{j\mathcal{I}K}, \vartheta_{j\mathcal{N}K}\}_{j\in S(\delta)}). \tag{21}$$

That is, the maximum value is achieved when all the families belong to the group whose ϑ is the highest. Therefore, by the definition of minimum, the optimal policy has to follow Equation (20). □

Unknown ϑ: When ϑ is unknown, the decision maker uses the data $\mathcal{Y}$ to learn about the true ϑ before choosing a policy δ. Therefore, the decision maker's risk is:

$$\mathcal{R}^{(\mathcal{N},\mathcal{I},city)}(\delta) = \mathbb{E}\left[\max_{\substack{\pi_{j\mathcal{I}}(\delta),\pi_{j\mathcal{N}}(\delta)\in\{0,1\},\text{ s.t} \\ \sum_j \left(\pi_{j\mathcal{N}}(\delta)+\pi_{j\mathcal{I}}(\delta)\right)=1}} \left\{ \sum_j \delta_j (\pi_{j\mathcal{I}}(\delta)\vartheta_{j\mathcal{I}K} + \pi_{j\mathcal{N}}(\delta))\vartheta_{j\mathcal{N}K}) \right\} |\mathcal{Y} \right], \tag{22}$$

where $\mathbb{E}[\cdot|\mathcal{Y}]$ denotes expectation with respect to the posterior distribution of ϑ given $\mathcal{Y}$.[38]

Optimal policy satisfies:

$$\boldsymbol{\delta}(\vartheta) = \underset{\delta,\ \text{s.t} \sum \delta_j = K}{\arg\min} \mathcal{R}^{(\mathcal{N},\mathcal{I},city)}(\vartheta, \delta),$$

and therefore:

$$\delta^{(\mathcal{N},\mathcal{I},city)}_{jK} = \underset{\delta \text{ s.t } \sum_j \delta_j = K}{\arg\min} \left\{ \mathbb{E}\left[\max(\{\vartheta_{j\mathcal{I}K}, \vartheta_{j\mathcal{N}K}\}_{j\in S(\delta)}) |\mathcal{Y}\right] \right\}, \tag{23}$$

where $S(\delta) = \{j : \delta_j = 1\}$ is the set of recommended cities.

Proof. Similar to when ϑ is unknown, the key part is observing that Equation (21) directly implies (23).

[38] Formmaly, risk is defined as $\int\int \mathcal{L}^{(\mathcal{N},\mathcal{I})}(\vartheta, \boldsymbol{\delta}(x)) f(x;\vartheta) dx dG(\vartheta)$. Standard arguments (e.g., in Wald (1950)) imply that minimizing it is equivalent to minimizing the posterior maximum risk in Equation (22).

L Restricted Choice Model

This section describes an extended model that restricts the sorting probabilities to better align with the spatial distribution in the data while maintaining compliance uncertainty. Let $\pi_{jg0} \in (0,1)$ be the share of group $g \in \{\mathcal{N}, \mathcal{I}\}$ individuals who live in city j in the absence of any policy such that $\sum_j (\pi_{j\mathcal{N}0} + \pi_{j\mathcal{I}0}) = 1$, and let $\pi_g(\delta) = (\pi_{jg}(\delta), ..., \pi_{Jg}(\delta))'$ be the location choice probability of group $g \in \{\mathcal{N}, \mathcal{I}\}$ given a policy δ. To rule out sorting probabilities that deviate from the status-quo sorting patterns, we consider only 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. The distance between $\pi(\delta)$ and π_0 is measured with the Total Variation distance function, which gives the largest absolute difference between the probability distributions across all 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 families' location choices follow the status quo distribution, while as $TV(\pi) \to 1$, all families belong to a single group and sort into a single location. We study the optimal policy under a hypothetical bound on the tendency of families to deviate from the status-quo shares: $TV(\pi) \leq a$, where $a \in [0,1]$ is the degree of compliance uncertainty. This restriction, together with the logical bound of $\pi_{jg}(\delta) \in [0,1]$ and of $\pi_{jg}(\delta) = 0$ if $\delta_j = 0$, implies that for every $g \in \{\mathcal{N}, \mathcal{I}\}$, location choices for city j with $\delta_j = 1$ is:

$$\pi_{jg}(\delta) \in [\max\{\tilde{\pi}^{\delta}_{jg0} - a, 0\}, \min\{\tilde{\pi}^{\delta}_{jg0} + a, 1\}], \quad \text{with} \quad \tilde{\pi}^{\delta}_{jg0} = \frac{\pi_{jg0}}{\sum_j (\pi_{j\mathcal{N}0} + \pi_{j\mathcal{I}0})\delta_j}, \tag{24}$$

where $\tilde{\pi}^{\delta}_{jg0}$ is the status quo shares normalized to sum to one across selected cities. These restrictions ensure that location choices approximately follow the status quo distribution while maintaining ambiguity regarding compliance ("Where do they go?") and families' group affiliation ("Who shows up?") governed by the parameter $a > 0$. The minimax decision-maker would like to choose a δ which is robust to the least favorable behavioral responses. For any δ, the maximum regret is:

$$\mathcal{L}^{max}_R(\vartheta, \delta) = \max_{\pi(\delta)} \mathcal{L}(\vartheta, \delta, \pi(\delta)) \text{ s.t Eqs (24)}, \ \sum_j (\pi_{j\mathcal{N}}(\delta) + \pi_{j\mathcal{I}}(\delta)) = 1, \tag{25}$$

where $\mathcal{L}(\vartheta, \delta, \pi(\delta))$ is defined in Equation (11), and δ is chosen to minimize: $\mathcal{R}^{\mathcal{N},\mathcal{I},city}_R(\delta) = \mathbb{E}[\mathcal{L}^{max}_R(\vartheta, \delta)|\mathcal{Y}]$, subject to $\sum_j \delta_j = K$. Similar to the policy in Equation (12), this decision rule ranks lists of size K places based on the expected maximum regret under the least favorable compliance. Unlike the unrestricted model, here the decision-maker assumes that there is a distribution of families across all recommended cities, ruling out the possibility that all 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 decision-maker faces more uncertainty regarding families' behavioral responses, and the optimal decision approaches the one reported in Equation (12).

Estimation: We estimate this decision rule with a bootstrap simulation. Given a value of $a \in [0, 1]$, we estimate the bootstrap average maximum risk: $\mathcal{R}^{*\mathcal{N},\mathcal{I},city}_{R}(\delta) = \mathbb{E}^*[\mathcal{L}^{max}(\vartheta, \delta)|\mathcal{Y}]$, where $\mathbb{E}^*$ is the expectation with respect to the S bootstrap draws from the posterior distribution of $\theta|\mathcal{Y}$, and in each bootstrap draw we solve Equation (25) by linear programing. The minimax policy is then: $\delta^{*\mathcal{N},\mathcal{I},city}_{K,R} = \arg\min_{\delta} \mathcal{R}^{*\mathcal{N},\mathcal{I},city}_{R}(\delta)$.

Resutls: Table L.1 reports the top 10 selected cities for $a = 0.01$ and $a = 0.9$. When location choices are restricted to align closely with the status quo distribution ($a = 0.001$), the selected cities provide low regret levels for native-born families, since they constitute the majority in each city in the status quo. In contrast, for $a = 0.9$, 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 with the oracle's first-best policy.

Table L.1: 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 natives (4)
Bat Yam	144.9	162.9	Bat Yam	144.9	162.9
Mevasseret Ziyyon	546.5	178.1	Haifa	222.1	289.4
Betar Illit	331.0	193.4	Rishon Leziyyon	267.4	260.7
Kokhav Ya'ir	501.0	224.1	Holon	288.4	292.2
Efrata	431.9	232.3	Karmi'el	214.6	307.4
Shomron	511.9	239.1	Qiryat Gat	298.7	257.1
Qiryat Eqron	556.9	288.8	Betar Illit	331.0	193.4
Bene Ayish	376.2	377.7	Ashdod	66.4	413.2
Qarne Shomeron	294.1	395.5	Efrata	431.9	232.3
Qiryat Ye'arim	310.3	600.1	Qiryat Motzkin	463.5	254.1

Note: This table reports the list of 10 selected cities by the restricted minimax ($\mathcal{N}/\mathcal{I}/city$) decision-maker depicted in Section L. Columns 1-2 report the posterior mean regret of the selected 10 cities when $a = 0.01$, and columns 3-4 report the posterior mean regret of the selected 10 cities when $a = 0.9$.

Microfoundation: Let $D_i \in \{1, ..., J\}$ be individual i's location choice to one of the J cities. We assume preferences follow the choice model: $D_i = \arg\max_{j\in\{1,...,J\}} U_{ij} - (a_{ij} + b_{ij}\delta_j)$, for every $g \in \{\mathcal{N}, \mathcal{I}\}$, where $U_i = (U_{i1}, ..., U_{iJ})$ is individual i's private valuation, whose pdf is $f(u|g, \delta)$. U_i doesn't have to follow a specific distribution and for $j \neq k$, U_{ij} and U_{ik} are allowed to be dependent. The share of families of group $g \in \{\mathcal{I}, \mathcal{N}\}$ who choose to move to location j given δ is:

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

where a_{gj} is set to match the status-quo sorting probabilities when $\delta_j = 0$ for all j, represented at $\delta = 0$: $\tilde{\pi}^{\delta}_{jg0} \equiv \pi_{jg}(0) = \int \mathbb{1}\{U_j - a_{gj} \geq U_k - a_{gk} \forall k\} f(u|g,0) du$. To see the equivalence between this model and the one in Section 10 let $\bar{b}_{gj}$ and $\underline{b}_{gj}$ be the lower- and upper-bounds of b_{ij} that satisfy: (1) If $\tilde{\pi}^{\delta}_{jg0} - a > 0$, then $\underline{b}_{gj}$ must solve:

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

while if $\tilde{\pi}^{\delta}_{jg0} - a \leq 0$, then $\underline{b}_{gj} \to -\infty$. (2) If $\tilde{\pi}^{\delta}_{jg0} + a < 1$, $\bar{b}_{gj}$ solves

$a = \displaystyle\int \mathbb{1}\{U_j - (a_{gj} + \bar{b}_{gj}\delta_j) \geq U_k - (a_{gk} + \bar{b}_{gl}\delta_k) \forall k\} f(u|g,\delta) du - \tilde{\pi}^{\delta}_{jg0}$, and $\bar{b}_{gj} \to \infty$ otherwise.

References

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

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

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

Kline, P., Rose, E. K., and Walters, C. R. (2024). A discrimination report card. *American Economic Review*, 114(8):2472–2525.

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

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

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