Immigrant Neighborhood Formation: Evidence from Canadian Cities*

Preliminary draft. Comments welcome.

Pierre-Loup Beauregard†, Vedran Razmilic‡, Pablo Valenzuela-Casasempere§

November 24, 2026

Abstract

Why do immigrant neighborhoods exist? We study how immigrant enclaves form in Montréal, Toronto, and Vancouver, using administrative data that follow where immigrants and native-born residents live every year. We first document that enclaves have changed substantially over the last thirty years: most of today's enclaves were not enclaves in 1991, and new enclaves have formed farther from city centers. To measure how different population segments value each neighborhood without restricting the dimensionality of local amenities, we adapt the PageRank algorithm to neighborhood-to-neighborhood moves. We microfound this approach in a dynamic model of neighborhood choice with forward-looking households, which gives the estimated valuations a clear indirect-utility interpretation. We show that those estimated indirect utility line up with neighborhood price premia from home sales and with access to local amenities. Immigrant and native-born valuations are positively but far from perfectly correlated. Using an instrumental variable approach, we estimate how each group responds to its neighbors. We find that every immigrant group values living near people from its own origin more than living near native-born residents, but the strength of this pull varies a lot across groups.

*We are grateful to Abdollah Farhoodi, Sam Gyetvay, Stephen Tino, and conference participants of the Immigration session at the 2026 CEA. This research was conducted at the University of British Columbia, a part of the Canadian Research Data Centre Network (CRDCN). This service is provided through the support of the Canada Foundation for Innovation, the Canadian Institutes of Health Research, the Social Sciences and Humanities Research Council, and Statistics Canada, as well as the support of UBC. The views expressed in this paper are our own and do not necessarily represent those of the CRDCN or its partners.

†Université de Montréal (pierre-loup.beauregard@umontreal.ca).

‡University of British Columbia (vedranrs@student.ubc.ca).

§Yale University (pablo.valenzuela-casasempere@yale.edu).

1 Introduction

Advanced economies are grappling with labor shortages and aging populations. These factors lead to a vicious cycle of rising wages, inflation exceeding central bank targets, and an increasing burden of healthcare expenditures on public finances. With economists' backing, policymakers have increasingly turned to immigration as a solution. As a result, immigrants now make up a significant and growing share of the population in developed countries.

One often overlooked key aspect of these south-north migration patterns is that immigrants tend to behave differently from native-born residents. These differences are particularly evident in how immigrants are distributed geographically. New immigrants often settle in existing co-ethnic communities, forming enclaves. The impacts of immigrant enclaves have been a major topic of study for economists, particularly those interested in immigrant integration in the labor market (Munshi, 2003; Damm, 2009). Researchers have also leveraged immigrant enclaves as instruments to obtain quasi-randomness to study the effect of immigration on wages (Altonji and Card, 1991; Dustmann et al., 2013). Yet, we know much less about where enclaves form, how they evolve, and what makes them persist.

This paper studies the formation of immigrant enclaves in Canada's three largest cities: Montréal, Toronto, and Vancouver. Using Canadian administrative data, we track the year-to-year residential moves of immigrants and Canadian-born tax filers. Our data combines tax files with immigration landing files, allowing us to see income from all sources, immigrants' country of origin, and, crucially, the place of residence every year. We first document that enclaves have changed substantially over the last thirty years. Most census tracts where immigrants made up at least half the population in 2021 were not enclaves in 1991, and newly formed enclaves lie farther from city centers than those that dissolved.

To understand what drives these patterns, we need to know how different groups value each neighborhood. We propose a simple approach to estimating neighborhood valuation in a setting where we do not restrict the dimension of local amenities. Instead of estimating the valuation of each possible amenity individually, our approach relies on individuals' location choices, which consider a bundle of neighborhood characteristics they value. This revealed-preference approach, popularized in labor economics by Sorkin (2018), is finding increasing application in urban economics (Bell et al., 2026, 2025; Fogel, 2021). We adapt the PageRank

algorithm (Page et al., 1999) to neighborhood-to-neighborhood moves and estimate a separate valuation of every neighborhood for each origin and income group.

We micro-found this approach in a dynamic model of neighborhood choice. Forward-looking households from different origins choose where to live, and each group values the endogenous composition of its neighborhood differently. Groups also spend different shares of their income on housing, so they react differently to price changes. Under the logit structure of the model and symmetric moving costs, bilateral flows between neighborhoods reveal each group's valuation of every neighborhood, which gives our estimates a clear indirect-utility interpretation. We do not need to solve for the equilibrium of the model to recover these valuations.

We then estimate how different segments of the population react to neighborhood population composition. Our framework incorporates endogenous feedback loops arising from the dynamic location decisions of households. Intuitively, demographic changes that occur today are not independent of those that will occur in the future; rather, changes in neighborhood demographics can affect neighborhood characteristics like property values and amenities, which in turn can influence demographics further, creating a cycle. We identify those response parameters to neighborhood demographics using an instrumental variable introduced by Caetano and Maheshri (2025), which exploits the fact that some residents stay in neighborhoods they chose based on information that is no longer relevant.

We find that immigrants and natives value the same neighborhoods quite differently. Our estimated valuations are positively correlated across groups, but native-born valuations explain at most a quarter of the variation in any immigrant group's valuations among richer households, and less than a tenth among poorer ones. Two forces drive these valuations: almost every group values the proximity to the city center, and each immigrant group is also pulled toward tracts where its own group represents a larger share. We benchmark our valuations to hedonic prices and local amenities, including access to schools, jobs, and health care. When we turn to the causal response to neighborhood composition, we find that every immigrant group values neighbors from its own origin more than native-born neighbors. The size of this premium varies a lot across groups. A 10 percentage point increase in the own-group share raises the valuation of Western households by about 8 log points and that of rich African and Middle Eastern households by about

7, compared with 1 to 3 log points for East and South Asian households. Responses to other immigrant groups are weaker and follow no clear pattern, while rich natives respond positively to every non-Western group and poor natives respond little to most of them. These own-group premia are the feedback loop at the heart of enclave formation: a neighborhood that gains co-nationals becomes more attractive to the next arrivals from the same origin.

Our work contributes to several existing pieces of literature. First, we add to the multidisciplinary literature on immigrant enclave formation. Enclaves substantially impact immigrant social and economic integration (Munshi, 2003; Edin et al., 2003; Damm, 2009; Beaman, 2012). It is, hence, not surprising to see immigrants of the same origin concentrate in some areas. While the old explanation of the "ports of entry" (Bartel, 1989) might not be as important as it used to be, the economic constraints and preferences are still important factors that explain immigrant segregation (Massey, 1990; Logan et al., 2002). Immigrant inflows also reshape the neighborhoods where immigrants settle. Saiz and Wachter (2011) show that housing values grow more slowly in U.S. neighborhoods where the immigrant share rises, consistent with native flight, and immigrant settlement has increasingly moved to the suburbs (Dawkins, 2009). Borrowing from Caetano and Maheshri (2025) in the context of ethnic segregation, we do not model those factors separately but rather estimate an indirect utility of each neighborhood, which includes all possible sources of segregation. We then recover how much of these valuations responds to neighbors from each origin group.

Second, we contribute to the literature studying neighborhood amenity values (Caetano and Maheshri, 2025; Bayer et al., 2016; Almagro and Domínguez-Iino, 2025; Ferreira and Wong, 2020; Bayer et al., 2004; Bayer and Timmins, 2005; Bayer et al., 2007; Albouy, 2016; Diamond, 2016). Our estimates differ from the existing literature in several ways. First, we estimate the neighborhood amenity value using tools from numerical linear algebra instead of relying on hedonic or structural approaches. Second, the object we estimate is slightly different from those papers: they usually estimate the amenity value of a location, whereas we estimate the full indirect utility of living in a location. The latter depends on amenities as well as labor market opportunities and cost of living. The third difference is the type of data required to estimate our amenity valuation. Most of the aforementioned articles use data on house purchases, whereas we use tax files that include both renters and homeowners. Taking information for all moving decisions, not only those

of homeowners, is an important aspect. Renters' moving decisions may differ from homeowners' because, when choosing their location, they don't internalize the property's future value in their decision.

Finally, our paper relates closely to Sorkin (2018) methodologically. Sorkin estimates the value of working at different firms using job-to-job transitions. We use the same revealed preference argument and an equivalent modified version of the PageRank algorithm (Page et al., 1999) to measure the value of living in a neighborhood using neighborhood-to-neighborhood moves. Contemporaneous work also applies this idea in urban economics, but evidence of the performance in the geographical context is scarce (Bell et al., 2026, 2025; Fogel, 2021). We contribute to applying this concept at a finer geographic level and mapping it to a standard dynamic model of location choice with forward-looking agents. This contextual shift enables us to uncover fresh insights and contributes to expanding the methodology's scope.

The rest of the paper is organized as follows. Section 2 describes the data. Section 3 documents how immigration and immigrant enclaves have evolved in Canadian cities. Section 4 presents the model, and Section 5 maps it to our estimation and reports the results. Section 6 concludes.

2 Data

This project combines administrative tax and immigration records with census data. The Longitudinal Administrative Databank (LAD) enables us to construct flows between neighborhoods, which are necessary for estimating our valuations. The Longitudinal Immigration Database (IMDB) identifies immigrants and their country of origin. The Census of Population provides housing expenditure shares and census-tract characteristics. Finally, we use property sales from British Columbia Assessment and Statistics Canada's Spatial Access Measures to benchmark our valuations (Appendices C and D).

2.1 Longitudinal Administrative Databank

The main component is Statistics Canada's Longitudinal Administrative Databank (LAD). The LAD is a longitudinal linkage of family tax files (T1FF) of a random 20% of Canadian workers. It allows us to follow taxpayers year to year from 1982 onward. The sampling of the dataset is such that if a worker is selected in a particular reference year, they will be included in any other later (or earlier) years in which they are

present in the T1FF database. Any individual who has a social insurance number and completed a T1 tax return or who received Canada Child Benefit can be selected into the LAD.[1]

The main variable of interest is the place of residence defined by the Census Tract (CT) of residence of the family. We identify the CT of residence from the full six-character postal code of the family. In urban areas, postal codes are very granular and usually identify an apartment building or a block face. We use the full six-character postal code and convert it to the 2021 census tract definition to avoid problems of statistical geographic units changing over time. We then track whether a household moved across neighborhoods by checking whether they filed taxes from a postal code that corresponds to a different CT in consecutive years.

The longitudinal nature of the LAD allows us to follow workers in time geographically but also to know their income every year. The primary income measure we use is Statistics Canada's definition of total pre-tax income. Total income is reported on line 15000 of the T1 tax form (line 150 before 2019) and refers to the sum of a tax filer's income for the Canada Revenue Agency's purposes. It includes employment income, rental income, self-employment net income, other income, pension income, Indian-exempt employment income, and benefits such as employment insurance, family and child benefits, social assistance, and tax credits, but excludes capital gains or losses. Due to confidentiality rules, amounts are rounded to the nearest $100.

In addition to income and place of residence, the LAD includes some limited characteristics of individuals and their families: the individual's age, gender, and marital status. T1 files also provide an array of amounts received as government transfers and tax credits from various sources. At the family level, the T1FF contains the family structure, its size, the number of children, and their respective dates of birth.

One advantage of using tax data is that it lets us follow individuals regardless of housing tenure. Most papers in the literature can follow only homeowners, as their mobility datasets are built from home purchase data (Caetano and Maheshri, 2025; Bayer et al., 2016). Tracking individuals beyond homeowners is key to our analysis, as renters comprise more than a third of all Canadian urban residents (Statistics Canada, 2023a).

We must note that while the LAD is a random sample of the tax-filer population, it is not representative of the whole population. However, Canadians have a very high incentive to file their tax returns even when

[1]This limits the longitudinal representativeness to individuals who have started filing income tax returns and their partners. According to Statistics Canada, the potentially sampled population covers around 75% of the official population estimates.

they have no employment income because of the large number of potential government transfers they can receive. This, however, means that cases where individuals lose all formal employment links, stop receiving transfers, or even lose a residential address will be overlooked.

2.2 Longitudinal Immigration Database

The Longitudinal Immigration Database (IMDB) is a comprehensive source of information on landed immigrants. The IMDB combines administrative files on immigrant admissions and non-permanent resident permits from *Immigration, Refugees and Citizenship Canada*. Landing records typically contain details collected by immigration officers during the application process, such as the immigrant's home country, age upon arrival, gender, education level, and immigrant category (e.g., economic programs, family reunion, refugees).

Landing record information is linked to individual tax records from the *Canada Revenue Agency* (CRA) in subsequent years. As for the LAD, we do not observe individuals who don't file tax forms. Again, individuals have a strong incentive to fill out their tax returns because they can potentially benefit from generous transfers.

We divide the immigrant sample into five origin groups: (*i*) other Western countries (Europe, the United States, Australia and New Zealand), (*ii*) Latin America and the Caribbean, (*iii*) Africa and the Middle East, (*iv*) East Asia, (*v*) South Asia and (*vi*) Oceania (excluding Australia and New Zealand).

2.3 Census

The Canadian Census of Population is conducted every five years and enumerates every household in the country. Participation is mandatory. All households answer a short-form questionnaire covering basic demographics, while a random sample of households answers a detailed long-form questionnaire covering, among other topics, income, immigration and place of birth, and housing: tenure, shelter costs, and, for owners, the value of the dwelling. The long-form sample covers one in five households in censuses up to 2006 and one in four households in 2016 and 2021.[2]

[2]In 2011, the mandatory long-form census questionnaire was replaced by the voluntary National Household Survey (NHS). Because participation was voluntary, only about 21% of all households responded. We treat the 2011 NHS as a regular census year throughout.

We use the Census of Population for two purposes: to measure housing expenditure shares by population group, which pin down the price coefficient in the model, and to measure housing prices at the census-tract level. We use the 1991 to 2021 censuses, which align with the five-year periods used in estimation.

2.4 Estimation Sample

Our estimation sample covers residents of the Montréal, Toronto, and Vancouver census metropolitan areas, aged 18 to 64, who file a tax return in two consecutive years. We split each origin group described in Section 2.2 and the native-born into poor and rich households, based on whether family income is below or above the city-year median, yielding 12 household groups. For each city, group, and census year $t \in \{1991, 1996, \ldots, 2021\}$, we pool the year-to-year transitions observed over the five tax years centered around t and build the bilateral matrix of within-city moves between census tracts, along with stay and exit counts. An exit is a move to an address outside the metropolitan area. Because the PageRank step requires every tract to both send and receive movers, we keep the largest strongly connected set of tracts in each matrix; dropped tracts receive no valuation for that group and period.

3 Stylized Facts

We present the stylized facts of our empirical setting and how they motivate our approach. We show that over the last 30 years, immigration in Canada has grown, with implications for traditional *immigrant* enclaves. The role of immigration flows, and how they interact with the way both locals and immigrants respond to compositional changes, is what motivates the overarching question of how immigrant enclaves are formed.

Fact 1: Immigration to Canada has grown steadily since the mid-1980s, with a marked shift in the origin composition of immigrants across cities.

Canada is one of the largest destination countries for migrants (United Nations, 2024). Figure 1 documents the total annual number of immigrants to Canada by region of origin between 1985 and 2021. Two trends stand out. First, aggregate immigration has grown substantially over the period, accelerating particularly after 2015. Second, the composition of immigrant flows has shifted dramatically. Europe, historically the dominant sending region, has declined in relative importance throughout the period. By contrast, im-

migration from South Asia (primarily India, Pakistan, and Sri Lanka) has grown rapidly, and by the end of the period South Asia is, together with Africa and the Middle East, one of the two largest source regions. Africa and the Middle East have similarly increased their share. Eastern Asia exhibited a distinct pattern, with a pronounced peak in the mid-1990s followed by a partial decline and stabilization.

These trends in immigration mask substantial differences across Canadian cities. Figure 2 shows that, for Toronto in the late 1980s, Europeans, South and Central Americans, and East Asians were the dominant groups among immigration inflows. Over a 35-year period, the immigration profile has shifted drastically, such that none of these 3 groups remain in the top 3 regions of origin. By 2015, South Asians were by far the most represented group, followed by immigrants from Africa and the Middle East.

Montréal displays a more balanced origin composition, with historically stronger flows from French-speaking countries reflecting Quebec's distinct immigration selection process (Beauregard et al., 2020). Throughout the period, immigration from Africa and the Middle East has dominated the total inflows. Vancouver presents perhaps the most striking city-specific pattern: a massive spike in Eastern Asian immigration in the late 1980s and 1990s, far exceeding that of any other origin group in any city, followed by a sharp decline. This spike reflects the wave of immigration from Hong Kong prior to its handover to China in 1997. However, in the late 2010s, this group of immigrants was surpassed by those from South Asia.

Together, these patterns suggest that the three largest Canadian cities are not simply scaled versions of the national trend but have each been shaped by distinct migration dynamics that are likely to leave different imprints on neighborhood composition and immigrant enclave formation.

Fact 2: Immigrant enclaves change over time and move further away from the city center.

As the immigrant population has grown since 1990, immigrants have settled in different neighborhoods than those of the previous immigration wave. Figures 3, 4, and 5 map the evolution of immigrant enclaves in Montréal, Toronto, and Vancouver between 1991 and 2021. Two important features stand out across all three cities. First, defining an immigrant enclave as a neighborhood where at least 50% of residents are immigrants, we find that most immigrant enclaves in 2021 were not enclaves in 1991 (tracts shown in light blue). This indicates that enclave formation has been both widespread and ongoing across Canadian cities. Second, while many 1991 enclaves persisted through 2021 (dark blue), several traditional immigrant

neighborhoods were no longer enclaves by 2021 (orange). Dissolved enclaves are geographically concentrated near the CBD, whereas newly formed enclaves tend to lie farther from the city center. This pattern is most pronounced in Toronto, where light-blue tracts span a wide suburban arc, while inner-city enclaves have been almost entirely replaced.

This pattern points to suburbanization of immigrant settlement over the last three decades. Figure 6 shows the distribution of immigrant neighborhoods by distance from the CBD in 1991 and 2021, pooling the three cities in our sample. As expected, immigrant neighborhoods in 2021 are located farther from city centers than they were in 1991. The average distance across census tracts rose from about 13.5 km in 1991 to 20 km in 2021, and the entire distribution shifted outward.

Taken together, these patterns suggest that enclave formation has been widespread and has shifted outward from city centers, consistent with the broader suburbanization of immigrant settlement documented in the literature (Dawkins, 2009). More generally, immigrant shares have evolved unevenly across neighborhoods, even within the same city.

4 A Dynamic Model of Neighborhood Choice

We build a dynamic model of a city's residential market that consists of (i) heterogeneous households making forward-looking location decisions and (ii) endogenous amenities determined by neighborhood composition. In the main text, households take housing prices as given. The model serves two purposes: it gives an indirect-utility interpretation to the valuations we recover from migration flows, and it defines the response parameters we estimate in Section 5.3. For these purposes, we only need the household side of the city. Appendix A.2 closes the model with a housing supply side, and Appendix A.3 defines its stationary equilibrium.

Notation. There are $J+1$ locations: J neighborhoods inside the city, indexed $j = 1, \ldots, J$, and an outside option $j = 0$. Households are classified into $K+1$ types: K immigrant groups indexed by country of origin and a native-born group N, each differing in their preferences over neighborhood composition. In the empirical application, we further split each group by income, so that a type is an origin–income pair

(Section 2.4). We define the population composition of neighborhood j at time t as the vector

$$M_{jt} \equiv \left[M_{jt}^1, \ldots, M_{jt}^K, M_{jt}^N\right]', \tag{1}$$

where M_{jt}^k is the number of type-k households in neighborhood j at time t, and we denote by $\mathbf{s}_{jt} \equiv M_{jt}/\sum_k M_{jt}^k$ the corresponding vector of population shares.

4.1 Endogenous Amenities

Following the literature (Ahlfeldt et al., 2015; Couture et al., 2023; Almagro and Domínguez-Iino, 2025; Caetano and Maheshri, 2025), amenities in neighborhood j are endogenously determined by the local population composition. A type-k household in neighborhood j at time t experiences amenity level

$$b_{jt}^k = G^k(M_{jt}), \tag{2}$$

where $G^k \colon \mathbb{R}_+^{K+1} \to \mathbb{R}_{++}$ is a continuous, bounded mapping. Indexing by k allows different groups to value the same neighborhood composition differently: an immigrant household may derive greater utility from co-ethnic neighbors than a native-born household does.

We adopt the following parametric specification:

$$\log b_{jt}^k = \mathbf{\Gamma}_b^k \mathbf{s}_{jt}, \tag{3}$$

where $\mathbf{\Gamma}_b^k$ is a $(K+1)$-dimensional row vector of type-k preferences for each group's presence in the neighborhood. Because shares sum to one, these preferences are identified only relative to one group, and we use the native-born share as the reference category. We abstract from spatial spillover preferences in which households also respond to the composition of nearby neighborhoods (Bagagli, 2025; Davis et al., 2025).

4.2 Housing Demand and Location Choice

4.2.1 Intra-Period Utility

After choosing a location, a type-k household has Cobb-Douglas preferences over a numeraire consumption good c_t and housing h_t, with expenditure share ϕ_k on the numeraire and $1-\phi_k$ on housing. The household maximizes

$$\max_{c,\,h} \; b_{jt}^{k} \left(\frac{c_t}{\phi_k}\right)^{\phi_k} \left(\frac{h_t}{1-\phi_k}\right)^{1-\phi_k} \quad \text{s.t.} \quad c_t + Q_{jt}\, h_t = w_{kt}, \tag{4}$$

where Q_{jt} is the price of housing in neighborhood j and w_{kt} is the income of type k at time t, assumed common across locations (perfect within-city labor market integration). The indirect utility is (Appendix A.1)

$$\bar{u}_t^k(j) \;=\; b_{jt}^{k}\, w_{kt}\, Q_{jt}^{\phi_k - 1}. \tag{5}$$

We define the *flow utility* as the logarithm of the indirect utility:

$$u_t^k(j) \;\equiv\; \log \bar{u}_t^k(j) \;=\; \log b_{jt}^{k} + \log w_{kt} + (\phi_k - 1)\log Q_{jt}. \tag{6}$$

A type-k household in neighborhood j demands $h_{jt}^{k} = (1-\phi_k)\, w_{kt}/Q_{jt}$ units of floor space, so it spends a constant share $1-\phi_k$ of its income on housing. The housing expenditure share is therefore measured directly from census microdata on household budgets for each group k, and the coefficient $\phi_k - 1 < 0$ on log housing prices is a known constant, not a free parameter (see Section 5.1).

4.2.2 Dynamic Location Choice

At the beginning of each period, households choose a location $j_t \in \{0, 1, \ldots, J\}$. Moving costs depend on the origin–destination pair and type:

$$MC^{k}(j_t,\, j_{t-1}) \;=\; \begin{cases} 0 & \text{if } j_t = j_{t-1}, \\ m_1^{k} & \text{if } j_t \neq j_{t-1},\; j_t \neq 0, \\ m_2^{k} & \text{if } j_t \neq j_{t-1},\; j_t = 0, \end{cases} \tag{7}$$

where $m_1^k > 0$ is the within-city moving cost and $m_2^k > 0$ is the cost of exiting to the outside option.[3] The flow payoff is

$$u_t^k(j_t,\, j_{t-1}) \;=\; u_t^k(j_t) - MC^k(j_t,\, j_{t-1}). \tag{8}$$

We collect the aggregate state in $\omega_t = (Q_t,\, b_t)$, where $Q_t = [Q_{1t}, \ldots, Q_{Jt}]'$ and $b_t = [b_{1t}, \ldots, b_{Jt}]'$. Household i of type k faces an i.i.d. Type I extreme value preference vector $\epsilon_{it} = [\epsilon_{i0t}, \ldots, \epsilon_{iJt}]'$ and solves

$$V_t^k(j_{t-1},\, \epsilon_{it}) \;=\; \max_{j_t} \Big[u_t^k(j_t,\, j_{t-1}) + \epsilon_{ij_tt} + \beta\, \mathbb{E}_t\big[V_{t+1}^k(j_t,\, \epsilon_{it+1}) \,\big|\, j_t \big] \Big]. \tag{9}$$

Markov structure. Location follows a first-order Markov process: the continuation value $EV^k(j) \equiv \mathbb{E}[V^k(j, \epsilon') \mid j]$ depends only on the destination j, not on the history $j_0, \ldots, j_{t-1}$. Any two households that choose the same destination $\tilde{j}$ therefore face identical continuation values regardless of their origin. Together with the logit structure, this property allows us to recover neighborhood values from observed choices without solving the dynamic program (Arcidiacono and Miller, 2011).

4.2.3 Location Choice Probabilities

Under the Type I extreme value assumption, the probability that a type-k household currently at j_{t-1} chooses j_t is

$$P_t^k(j_t \mid j_{t-1}) \;=\; \frac{\exp\big\{ u_t^k(j_t,\, j_{t-1}) + \beta\, EV^k(j_t) \big\}}{\displaystyle\sum_{j'=0}^{J} \exp\big\{ u_t^k(j',\, j_{t-1}) + \beta\, EV^k(j') \big\}}. \tag{10}$$

From the model to the estimation. Two features of the model carry over to the estimation. First, with logit shocks and symmetric within-city moving costs, the ratio of bilateral flows between two neighborhoods depends only on how each group values the two neighborhoods, which is what the PageRank step exploits (Section 5.2). Second, Cobb-Douglas preferences pin down the coefficient on housing prices, so that we can net prices out of the valuations before projecting them onto neighborhood composition (Section 5.3). We do not need to solve for the equilibrium of the model, nor to estimate moving costs, to recover either object. Appendix B.2 shows that the same flows also identify moving costs, which are needed to simulate the model.

[3] Moving costs are constant within the city. This assumption is standard in the dynamic location-choice literature (Almagro and Domínguez-Iino, 2025) and is needed for the symmetric cancellation in the PageRank step (Section 5.2).

5 Estimation and Results

The model is estimated in three steps, exploiting the logit structure of location choices to avoid solving the full dynamic program. Each subsection presents one step of the procedure (the estimating equation, the identifying assumptions, and the implementation) and then reports the corresponding estimates. Section 5.1 recovers housing expenditure shares from census microdata. Section 5.2 recovers neighborhood valuations via PageRank. Section 5.3 projects the neighborhood valuations onto neighborhood composition to estimate preferences for neighbors. Because the valuations bundle the flow utility with the discounted continuation value, none of these steps requires us to calibrate the discount factor β or to estimate moving costs.

5.1 Housing Expenditure Shares

Estimating equation. Under the Cobb-Douglas preferences of Section 4.2.1, a type-k household spends a constant share $1-\phi_k$ of its income on housing, so the coefficient on log housing prices is $\phi_k - 1$ and does not need to be instrumented. We compute ϕ_k directly from census microdata as

$$\hat{\phi}_{k,c,\tau} \;=\; 1 - \text{med}_{i \in k,c,\tau}\left\{\frac{\text{shelter cost}_i}{\text{household income}_i}\right\}, \tag{11}$$

for each group k, city c, and census year τ. We exclude households reporting zero shelter costs or zero income; no household in the sample reports negative values.

The model assumes ϕ_k is constant over time and across locations within a city. We find evidence for this assumption in the data, with minimal difference in $\hat{\phi}_{k,c,\tau}$ across census years. Therefore, we average across years to obtain one share per group and city, $\hat{\phi}_k$.

Table 1 reports the estimated housing expenditure shares. Three patterns stand out. First, in almost every city and income group, immigrants spend a larger share of their income on housing than natives. The few exceptions are immigrants from Europe and the United States in Vancouver and rich ones in Toronto, as well as poor immigrants from Oceania in Toronto, whose shares are within one percentage point of natives'. Second, the share spent on housing is smaller for rich households in every group and city, in line with the non-homotheticity of housing consumption (Couture et al., 2023). Finally, there is substantial cross-city variation in housing expenditures. For example, poor South Asian households spend around 31.7% of their

yearly income on housing costs in Montréal, compared with 41.2% in Vancouver.

5.2 PageRank and Neighborhood Valuations

The second step recovers type-specific values $\widetilde{V}^k(j)$ for each group k, neighborhood j, and period t from the bilateral migration flow matrix, without solving the Bellman equation.

Flow-relevant values. Define the flow-relevant value as

$$\widetilde{V}^k(j) \equiv \frac{e^{u^k(j) + \beta\, EV^k(j)}}{M_j^k}, \tag{12}$$

where $u^k(j)$ is the flow utility, $EV^k(j)$ is the expected continuation value, and M_j^k is the mass of type-k residents. The object $\widetilde{V}^k(j)$ bundles current and future neighborhood value, normalized by population.

PageRank algorithm. We adapt Sorkin's (2018) method to average out idiosyncratic shocks and extract the systematic pattern of mobility. This method is based on the PageRank algorithm, originally developed to rank websites by their importance in the web graph (Page et al., 1999). Intuitively, highly valued neighborhoods have more inflows (links) from other highly valued neighborhoods. To our knowledge, the only applications of this idea in spatial economics are Fogel (2021) and Bell et al. (2026, 2025).

We adapt the PageRank algorithm to rank neighborhoods. We compute the ranking using individuals moving within the city. To simplify notation, the equations below omit the time-period subscript. However, we compute the valuation for each specific group and time period. Under the logit structure with symmetric within-city moving costs (Section 4.2.2), the bilateral flow ratio satisfies $\Upsilon^k_{h\to j}/\Upsilon^k_{j\to h} = \widetilde{V}^k(j)/\widetilde{V}^k(h) \cdot [1 + O(J^{-1})]$ (see Appendix B.1 for the derivation). Cross-multiplying and summing over origins yields the PageRank fixed-point equation:

$$\widetilde{V}^k(j) = \frac{\sum_{h\in\mathcal{J}} \Upsilon^k_{h\to j}\, \widetilde{V}^k(h)}{\sum_{h\in\mathcal{J}} \Upsilon^k_{j\to h}}, \qquad \forall\, j \in \mathcal{J}, \tag{13}$$

where $\mathcal{J} = \{1, \ldots, J\}$ is the set of within-city neighborhoods and $\Upsilon^k_{h\to j}$ is the number of type-k individuals moving from neighborhood h to neighborhood j.

This object is very intuitive in terms of valuation. A neighborhood valuation is high if, once people move into the neighborhood, they do not move out (low out-migration), or if people moving into the neighborhood come from other highly valued neighborhoods (Sorkin, 2018). Analogously, a neighborhood valuation is low if it has high out-migration and low in-migration from other highly valued neighborhoods.

Equation (13) can be rewritten in matrix form. Let $\boldsymbol{\Upsilon}^k$ be the $J \times J$ matrix of moves between neighborhoods, whose entry in row j and column h is $\Upsilon^k_{h \to j}$, and let $\mathbf{S}^k$ be the diagonal matrix of total out-migration, with $S^k_{jj} = \sum_{h \in \mathcal{J}} \Upsilon^k_{j \to h}$. Finally, let $\widetilde{\mathbf{V}}^k$ be the vector of flow-relevant valuations for each neighborhood. This gives the following matrix form:

$$\widetilde{\mathbf{V}}^k = (\mathbf{S}^k)^{-1} \boldsymbol{\Upsilon}^k \widetilde{\mathbf{V}}^k. \tag{14}$$

Sorkin (2018) shows that equation (14) is solvable, up to a multiplicative factor, if the set of neighborhoods is strongly connected. A strongly connected neighborhood receives migrants from other neighborhoods and has households migrating to other neighborhoods. This condition is easily met when individuals move across a limited set of neighborhoods, usually fewer than in firm-to-firm transitions.

Implementation. For each group k, city, and period t: (i) construct the bilateral flow matrix $\hat{\boldsymbol{\Upsilon}}^k$ from the longitudinal tax records, restricted to the largest strongly connected set of tracts (Section 2.4); (ii) solve (14) for the eigenvector of $(\hat{\mathbf{S}}^k)^{-1} \hat{\boldsymbol{\Upsilon}}^k$ associated with the eigenvalue one; (iii) normalize $\sum_j \widetilde{V}^k(j) = J$. Step (ii) is equivalent to computing the stationary distribution of the Markov chain of within-city movers and dividing it by each tract's total out-migration (Appendix B.1.7). Tracts outside the strongly connected set receive no valuation for that group and period.

Because the flow-ratio condition holds pairwise, restriction to the strongly connected set does not affect the relative valuation of the tract within the set. Excluded tracts are those with too few movers from group k. Their exclusion affects only the period normalization, which scales all valuations of a group in a given city-period by a common factor. In Section 5.3, our regression is estimated separately for each group and that common factor is absorbed by the fixed effects. The rest of the analysis is based on within-group comparisons of valuation, which are unaffected by the scale normalization.

Results. Figures 7, 8, and 9 map the estimated 2021 valuations for natives and the two largest immigrant groups in the metropolitan areas of Toronto, Montréal, and Vancouver. Three patterns emerge. First, we find that in all three metros, immigrants rank neighborhoods closer to the city center higher and outer neighborhoods lower than natives with the same income. Immigrant groups value central neighborhoods more: on the island of Montréal around downtown, in downtown Vancouver and the city's west side, and throughout much of downtown Toronto. Within each origin group, income shifts valuations toward the center. Rich immigrants value the city center more uniformly than poor immigrants, whose high valuations are more scattered across the suburbs. Rich natives in Montréal and Vancouver favor central areas that poor natives do not, although rich natives in Toronto rank downtown as low as poor natives do. However, income does not close the gap between immigrants and natives. Even among poor households, immigrants value the core highly while natives do not.

The second pattern we found is that immigrants and natives also differ in how spread out they are across the metropolitan area. Natives have valuations in nearly every neighborhood. The few tracts without one are the same for rich and poor natives and are largely uninhabited. Immigrants, by contrast, are absent from substantial parts of each metropolitan area, and where those gaps fall differs by city. Finally, the maps show that immigrants value living in enclaves. Immigrants' highest valuations form a few contiguous zones (i.e., enclaves), whereas natives' highest valuations are spread across many parts of the metropolitan area. This concentration is strongest among rich immigrants; for instance, the highest-valued neighborhoods for rich African and Middle Eastern immigrants in Montréal cluster tightly around downtown. Immigrant valuation maps are also patchier, with adjacent neighborhoods often ranked very differently, while native maps vary more smoothly across large areas.

Tables 2 and 3 report how valuations correlate across groups (Panel A) and with observable neighborhood characteristics (Panel B) for rich and poor households, respectively. We find that immigrant and native valuations are positively correlated, but far from perfectly. Among rich households, the correlation ranges from 0.30 for East Asians to 0.49 for Africans and Middle Easterners, implying that native valuation explains at most 24 percent of the within-city variation in the valuation of any immigrant group. Among poor households, the correlations fall to 0.21–0.29, and this share drops to at most 9 percent. Panel B shows how

immigrants and natives value neighborhood characteristics differently. The valuation of every immigrant group rises with the share of its own members in the tract, with correlations between 0.08 and 0.24. In contrast, the native valuation falls with the native share (-0.57 among the rich and -0.30 among the poor). Because $\widetilde{V}^k$ is normalized by the group's population in the tract (equation (12)), part of this negative correlation may be mechanical. All groups, however, place a higher valuation on tracts closer to downtown. Excluding poor natives, whose valuation is essentially unrelated to distance to the city center, the correlation with log distance to the CBD ranges from -0.43 to -0.20. Finally, valuations are, if anything, negatively correlated with log rent and log median income, with correlations between -0.28 and -0.01. The valuations do not simply reflect how expensive a neighborhood is. Together, these results suggest that the valuations capture a pull toward the city center that almost all groups share, and a pull toward co-nationals that is specific to each immigrant group.

Comparison against hedonic valuation. We compare our neighborhood valuation against the corresponding neighborhood valuation a hedonic model would imply. To do so, we use British Columbia Assessment (BCA) data, which records property characteristics and information about the last three recorded sales. We run a hedonic regression of log home prices on property characteristics and time-varying neighborhood (census tract) fixed effects, and convert these fixed effects into within-year percentiles.[4] What we call the hedonic valuation is, therefore, the price premium common to all properties in a census tract and period. One caveat of our dataset is that we do not have any information about the buyer, limiting the across-immigrant-group analysis. We therefore re-estimate the PageRank neighborhood valuation for the city of Vancouver, aggregating all groups into one category to match BCA's data. More details on this exercise appear in Appendix C.

Figure C.1 shows the relationship between the two measures of the neighborhood valuations. Across the six years, we find a strong positive relationship between these two measures, although the slope is below unity, suggesting heterogeneity between them. This is similar to what has been documented in labor economics, where the value of working at a firm and the pay it offers do not track one-to-one (Sorkin, 2018). The valuation estimated using migration flows does a good job of predicting the variation in the hedonic

[4]We choose the year partition to match our PageRank years. Our BCA agreement includes property sales up until 2021, so the last census year aggregate sales occur between 2018 and 2021.

valuation. The R-squared values range from 0.066 in 2001 to 0.400 in 1991 and exceed 0.2 in three of the six years.

Benchmarking against observable neighborhood amenities. Statistics Canada released spatial access measures for each dissemination block in 2021 (Statistics Canada, 2023b). These measures capture access, by transit and by walking, to education, employment, health care, and other amenities, which we aggregate to the census tract level using population-weighted averages. More details on how we construct these amenity measures can be found in Appendix D.1.

We benchmark our neighborhood valuations against these amenity indexes. Since every amenity, and our valuations, are in different units, we compare the percentiles in the corresponding city distribution. Figure D.1 presents the relationships between Statistics Canada's amenity measures and our neighborhood valuations for the pooled sample in the city of Vancouver. Every row corresponds to a different regression. We find two overarching patterns. First, the estimated valuations correlate positively with better access to these amenities. Walking access to primary and secondary schools has the strongest relationship with the valuations, with a coefficient of almost 0.5. Access to almost every amenity is positively and significantly related to the valuations, which lends credibility to our measure. Second, with the exception of access to primary and secondary schools, access by transit matters more than walking access, as the corresponding coefficients are larger.

Figure D.2 repeats the exercise separately for each immigrant-income group. Each panel corresponds to a different amenity. We find that each amenity is valued differently by each immigrant-income group, with large heterogeneity in valuations. For example, the coefficient on employment access for rich South Asian and rich Western (Europe and USA) households is about double that of rich natives. Poor natives stand out: their valuation is negatively related to transit access to every amenity, consistent with their valuation being essentially unrelated to distance to the city center.

5.3 Preferences for Neighborhood Composition

With the PageRank valuations in hand, we project them onto observables to recover preferences for neighborhood composition. Let $v_{jt}^k \equiv \log \hat{\widehat{V}}_{jt}^k$ denote the estimated log valuation of neighborhood j for type k in

period t. Because v_{jt}^k bundles the flow utility with the discounted continuation value (equation (12)), we specify it directly as

$$v_{jt}^k \;=\; \delta_j^k + \delta_t^k + (\phi_k - 1)\,\log Q_{jt} + \mathbf{\Gamma}_b^k\,\mathbf{s}_{jt} + \xi_{jt}^k, \tag{15}$$

where ϕ_k is the Cobb-Douglas housing expenditure share estimated in Section 5.1, Q_{jt} is the census-tract housing price measured in the Census of Population (Section 2.3), $\mathbf{s}_{jt}$ is the vector of group population shares, $\mathbf{\Gamma}_b^k$ captures type-k responses to neighborhood composition, δ_j^k and δ_t^k are neighborhood–type and period–type fixed effects, and ξ_{jt}^k is a time-varying unobservable. The fixed effects absorb time-invariant sorting patterns and group-level income trends (w_{kt}); ξ_{jt}^k captures transient neighborhood-specific shocks to amenities, preferences, or expectations.

Because ϕ_k is known, the coefficient on log housing prices is a fixed constant. We move the housing price term to the left-hand side and define the *price-adjusted valuation*:

$$\tilde{v}_{jt}^k \;\equiv\; v_{jt}^k - (\phi_k - 1)\,\log Q_{jt} \;=\; \delta_j^k + \delta_t^k + \mathbf{\Gamma}_b^k\,\mathbf{s}_{jt} + \xi_{jt}^k. \tag{16}$$

This transformation eliminates $\log Q_{jt}$ as a regressor, removing one source of endogeneity and the need to instrument for housing prices in the demand-side estimation. The adjustment removes the effect of current prices only; anticipated changes in future prices remain in the continuation value and are absorbed by $\mathbf{\Gamma}_b^k$ and ξ_{jt}^k.

Endogeneity. The remaining endogenous regressor is neighborhood composition, which we measure with the population shares of each origin group, $s_{jt}^{k'} = M_{jt}^{k'} / \sum_{k''} M_{jt}^{k''}$, omitting the native-born share as the reference category. Two forces bias OLS estimates of Γ_b^k. First, groups with similar tastes sort into the same neighborhoods, so a neighborhood with a high unobserved amenity ξ_{jt}^k also hosts many type-k households. Second, $\mathbf{s}_{jt}$ is itself the aggregate outcome of the location choices whose value we are measuring, so a shock to ξ_{jt}^k mechanically moves $\mathbf{s}_{jt}$ in the same period. Neighborhood fixed effects absorb time-invariant sorting but not time-varying amenity shocks. Both forces bias $\hat{\Gamma}_b^k$ away from zero, especially for the own-group coefficient.

Identification. We follow Caetano and Maheshri (2025) and exploit the frictions in the model to build an instrument. Because moving is costly and households cannot perfectly forecast how neighborhoods will evolve, some residents of j at any date chose it based on expectations that later proved wrong. Having become mismatched, many of them stay because their loss of utility is smaller than the moving cost m_1^k. These households still contribute to the current composition $\mathbf{s}_{jt}$, which gives relevance. The information that originally drew them to j is no longer relevant to the households choosing locations at t, which gives validity. Formally, past composition must carry no information about current unobserved amenities beyond what is already embedded in more recent valuations:

$$\mathbb{E}\big[\xi_{jt}^k \,\big|\, \tilde{\mathbf{v}}_{j,t-T},\, \mathbf{s}_{j,t-(T+1)},\, \delta_j^k,\, \delta_t^k\big] = \mathbb{E}\big[\xi_{jt}^k \,\big|\, \tilde{\mathbf{v}}_{j,t-T},\, \delta_j^k,\, \delta_t^k\big], \tag{17}$$

where $\tilde{\mathbf{v}}_{j,t-T}$ stacks the price-adjusted valuations of all groups in neighborhood j at $t-T$. Any information from $t-(T+1)$ or earlier that households use at t to sort on amenities must therefore already have been used by households that sorted at $t-T$, and so be reflected in their valuations. The assumption does not impose myopia; it only rules out households at t using old information in a more sophisticated way than households at $t-T$. Unlike a shift-share design, the lagged shares need to be exogenous only conditional on lagged valuations, not unconditionally.

Relevance requires that $\mathbf{s}_{j,t-(T+1)}$ predict $\mathbf{s}_{jt}$ conditional on $\tilde{\mathbf{v}}_{j,t-T}$. This holds only if both frictions are present. Without moving costs, mismatched households would leave immediately. Without uncertainty, no household would become mismatched. We test relevance directly below, and the first stage is strong.

Implementation. For each group k, we estimate

$$\tilde{v}_{jt}^k = \delta_{c(j)}^k + \delta_t^k + \Gamma_b^k\, \mathbf{s}_{jt} + \rho^k\, s_{jt}^R + \boldsymbol{\gamma}^k\, \tilde{\mathbf{v}}_{j,t-T} + \varepsilon_{jt}^k \tag{18}$$

by two-stage least squares, using the lagged shares $\mathbf{s}_{j,t-(T+1)}$ as excluded instruments for $\mathbf{s}_{jt}$, where s_{jt}^R is the share of rich households in tract j. Periods are five-year census intervals and we set T to one period. The controls are therefore all groups' valuations five years earlier, and the instruments are the neighborhood composition ten years earlier. The lagged valuations absorb persistent differences across tracts, so we replace

the tract fixed effects of equation (16) with group-city fixed effects $\delta^k_{c(j)}$ and pool all cities together. Because each equation has several endogenous shares, we report the Kleibergen–Paap rk Wald F statistic. Standard errors are clustered at the census-tract level.

Interpretation. Because the dependent variable bundles flow utility and continuation value, $\mathbf{\Gamma}^k_b$ measures the total response of type-k valuations to neighborhood composition, through current amenities and through expected future composition, prices, and amenities. This is the same object as the response parameters in Caetano and Maheshri (2025). It is not a pure taste parameter: it captures every reason type-k households respond to neighborhood composition, including taste-based homophily, statistical inference about other amenities such as schools or co-ethnic services, and supply-side steering that limits the neighborhoods available to some groups. We do not attempt to separate these channels in this paper.

Results. Table 4 reports the estimated composition parameters $\mathbf{\Gamma}^k_b$ for poor and rich households of each group k, pooling the three cities. The native-born share is omitted, so each coefficient measures how the valuation of group k changes when households of a given origin group replace native-born neighbors, holding the share of rich households fixed. The coefficients are interpreted as a combination of both pure taste and changes in the expectation of the neighborhood, because they bundle the direct value of neighbors with the expected evolution of the neighborhood that a change in composition creates. The first stage is strong, with a Kleibergen–Paap F-statistic of about 240 in every column. We find that every immigrant group values neighbors from its own group more than native-born neighbors. All ten own-group coefficients are positive, and nine are statistically significant at the 1 percent level. The response is largest for Western households. A 10 percentage point increase in the Western share raises their valuation by about 8 log points, among both the poor and the rich. The same increase in the own-group share raises the valuation of South Asian households by 2–3 log points and that of poor East Asian households by only 1 log point. Native-born households respond to Western neighbors in the opposite direction, and a 10 percentage point increase in the Western share lowers their valuation by 4.5 log points among the poor and by 3.3 among the rich. Together, these estimates show that every immigrant group gains from living near its own group, which is the feedback through which composition can reinforce itself in the model, leading to immigrant enclaves.

Three further patterns stand out beyond the own-group premium. First, income shapes the strength of the own-group response, but not in the same direction for every group. A 10 percentage point increase in the own-group share raises the valuation of rich African and Middle Eastern households by 6.7 log points, twice the response of poor households from the same region. Rich East Asian households also respond more than twice as strongly as poor ones, while Latin American households show the reverse pattern: 5.9 log points for the poor and 3.4 for the rich. Second, responses to other immigrant groups are weaker and less systematic. Only 10 of the 40 cross-group coefficients among immigrant households are statistically significant at the 5 percent level, and they split evenly between positive and negative signs. The largest are the responses of poor African and Middle Eastern households and of rich East Asian households to the Latin American share, which is also the least precisely estimated row in the table. Third, beyond their negative response to Western neighbors, native-born households show a sharp income gradient. Rich natives respond positively to the share of every non-Western origin group in the table, with their valuation rising by 2.6–4.1 log points for a 10 percentage point increase. Poor natives, by contrast, respond to these groups weakly or not at all, except for a positive response to the South Asian share. Because the specification conditions on the share of rich households, these responses do not proxy for the overall income mix of the neighborhood, although they may still reflect the incomes of the particular groups that move in. Overall, we find heterogeneous responses to both own-group and out-of-group shares, which together shape the distribution of urban residents in Canadian cities.

6 Conclusion

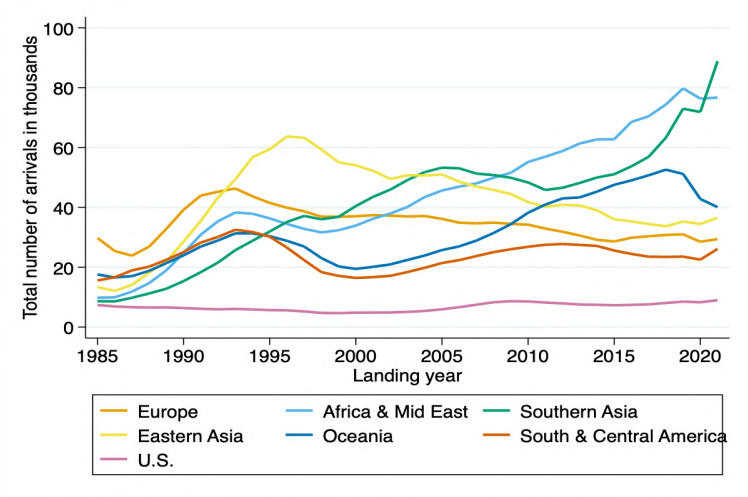
This paper studies the formation of immigrant enclaves in Canada's three largest cities. Over the last thirty years, enclaves have formed, dissolved, and moved away from city centers. Using neighborhood-to-neighborhood moves in administrative data, we recover how each origin and income group values every neighborhood, and we ground these valuations in a dynamic model of neighborhood choice. The valuations differ substantially across groups, line up with hedonic price premia and with access to local amenities, and reveal a shared pull toward the city center along with a group-specific pull toward co-nationals. Instrumenting current neighborhood composition with its lagged values, we find that every immigrant group values

neighbors from its own origin more than native-born neighbors, with large differences in the strength of this response across groups.

These own-group responses imply that neighborhood composition can reinforce itself: a neighborhood that attracts co-nationals today becomes more attractive to the next arrivals from the same origin. In ongoing work, we combine these estimates with the housing supply side and equilibrium of Appendix A.3 and the moving costs of Appendix B.2 to simulate how immigration levels, the income profile of immigrants, and their origins shape where enclaves form and how long they persist.

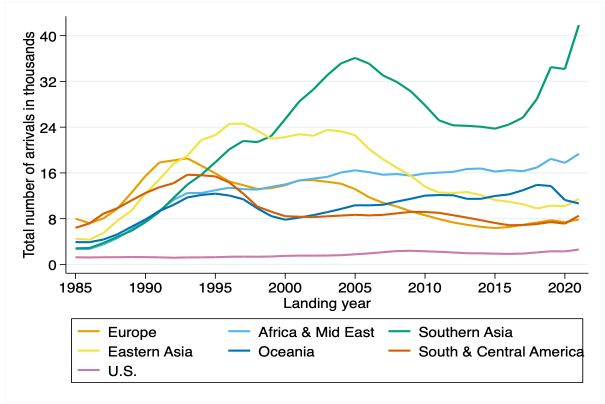
7 Figures

Figure 1: Immigration Time Series in Canada

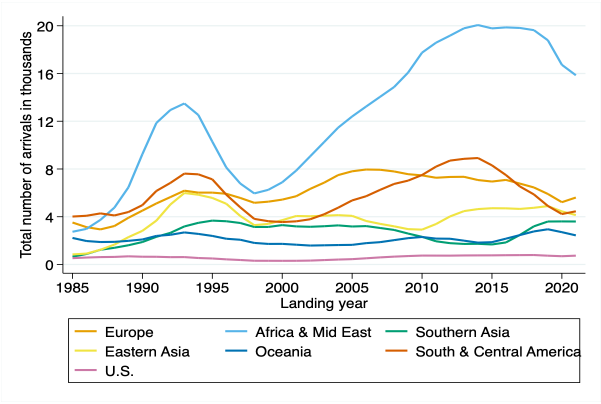


Notes: The figure presents the total number of immigrants by region of origin in Canada from 1985 to 2021.

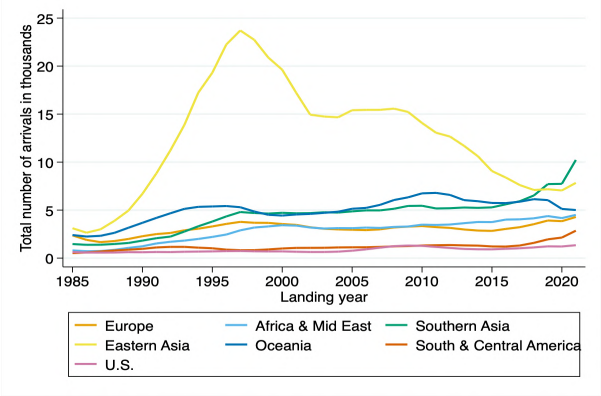
Figure 2: Immigration Time Series in Major Canadian Cities



(a) Toronto



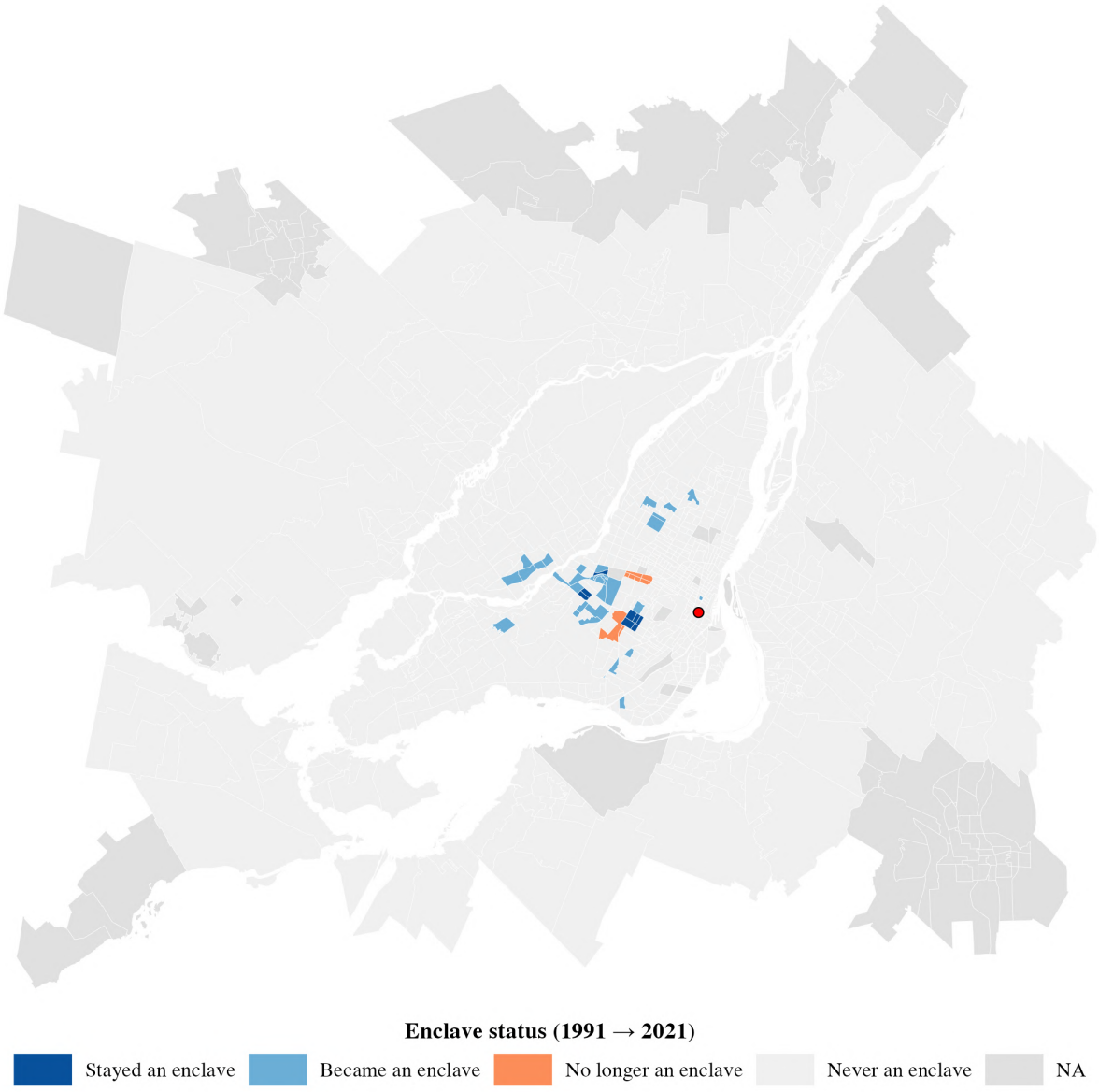
(b) Montreal



(c) Vancouver

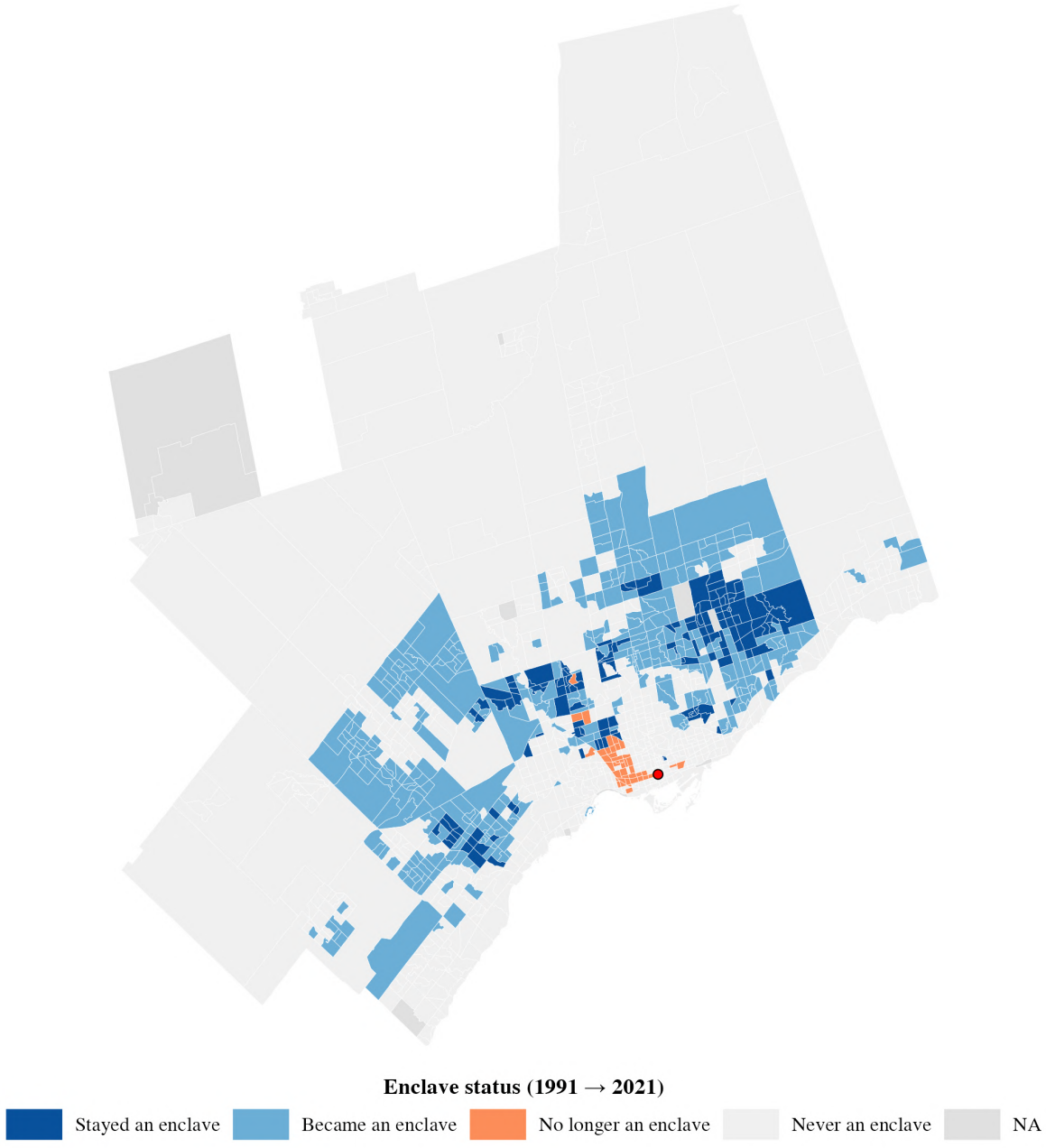
Notes: The figure presents the total number of immigrants arriving in Canada's three largest cities (Toronto, Montreal, and Vancouver) from 1985 to 2021. The top-left panel presents the graph for Toronto, the top-right panel for Montreal, and the bottom panel for Vancouver.

Figure 3: Evolution of Immigrant Enclaves in Montréal



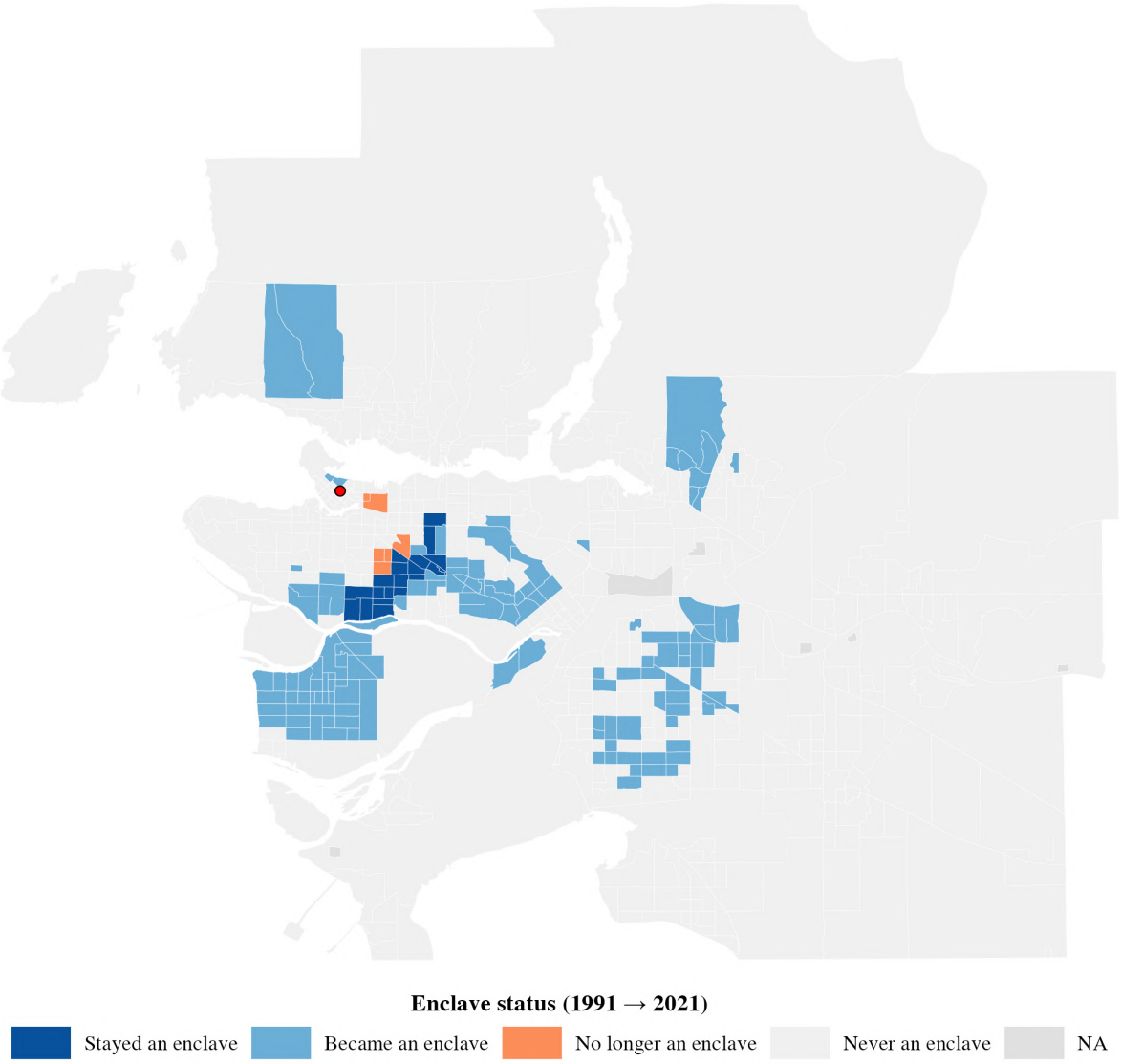
Notes: The maps document the evolution of immigrant enclaves across census tracts in Montréal between 1991 and 2021. A census tract is classified as an enclave if immigrants comprise at least 50% of its resident population. Dark blue tracts were enclaves in both 1991 and 2021. Light blue tracts were not enclaves in 1991 but had crossed the 50% threshold by 2021. Orange tracts were enclaves in 1991 but fell below the threshold by 2021. Light grey tracts do not meet the enclave threshold in either year. The red dot indicates the location of the Central Business District. The 1991 immigrant share is projected onto 2021 census geographies using the Canadian Longitudinal Tract Database (Allen and Taylor, 2018).
Sources: 1991 and 2021 Census Profiles Files, retrieved using Canadian Census Analyser.

Figure 4: Evolution of Immigrant Enclaves in Toronto



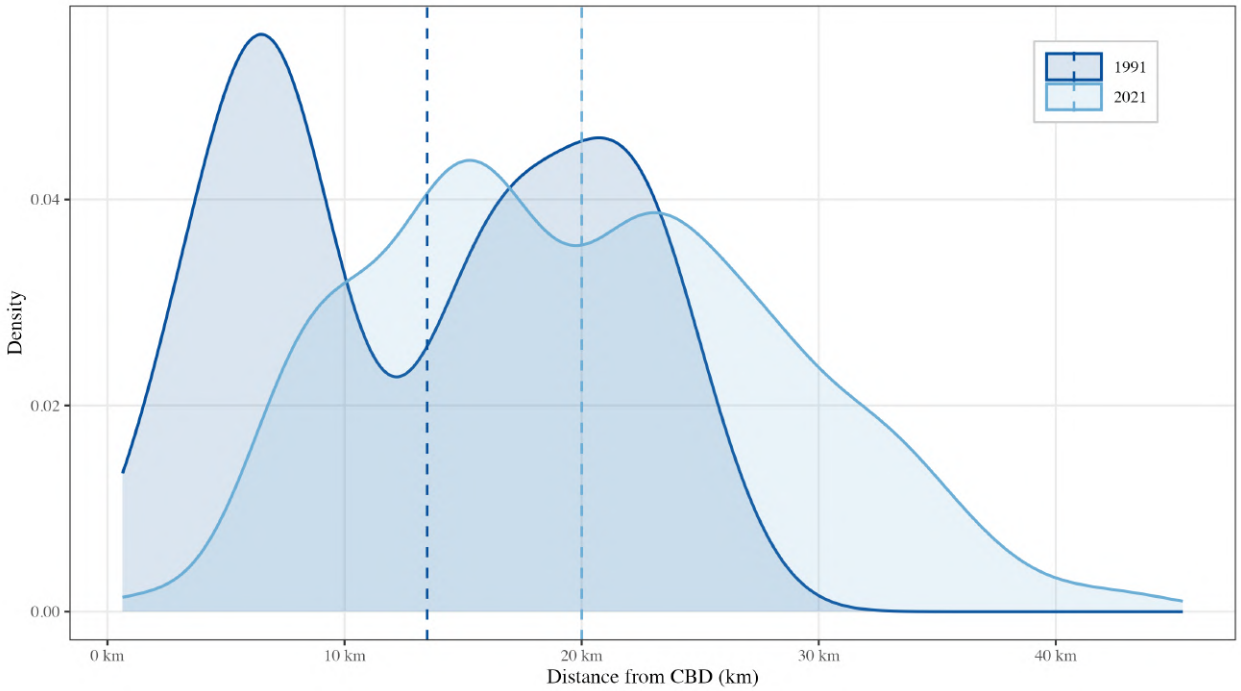
Notes: The maps document the evolution of immigrant enclaves across census tracts in Toronto between 1991 and 2021. A census tract is classified as an enclave if immigrants comprise at least 50% of its resident population. Dark blue tracts were enclaves in both 1991 and 2021. Light blue tracts were not enclaves in 1991 but had crossed the 50% threshold by 2021. Orange tracts were enclaves in 1991 but fell below the threshold by 2021. Light grey tracts do not meet the enclave threshold in either year. The red dot indicates the location of the Central Business District. The 1991 immigrant share is projected onto 2021 census geographies using the Canadian Longitudinal Tract Database (Allen and Taylor, 2018).
Sources: 1991 and 2021 Census Profiles Files, retrieved using Canadian Census Analyser.

Figure 5: Evolution of Immigrant Enclaves in Vancouver



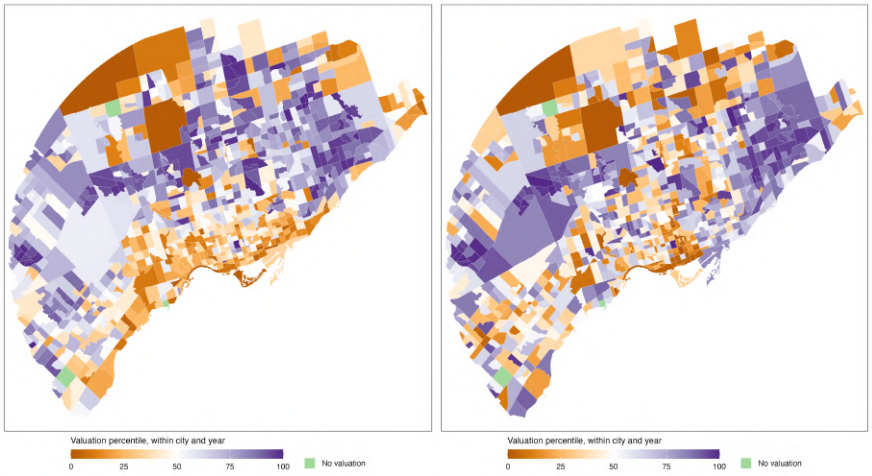
Notes: The maps document the evolution of immigrant enclaves across census tracts in Vancouver between 1991 and 2021. A census tract is classified as an enclave if immigrants comprise at least 50% of its resident population. Dark blue tracts were enclaves in both 1991 and 2021. Light blue tracts were not enclaves in 1991 but had crossed the 50% threshold by 2021. Orange tracts were enclaves in 1991 but fell below the threshold by 2021. Light grey tracts do not meet the enclave threshold in either year. The red dot indicates the location of the Central Business District. The 1991 immigrant share is projected onto 2021 census geographies using the Canadian Longitudinal Tract Database (Allen and Taylor, 2018).
Sources: 1991 and 2021 Census Profiles Files, retrieved using Canadian Census Analyser.

Figure 6: Evolution of Immigrant Enclaves Relative to Central Business Districts



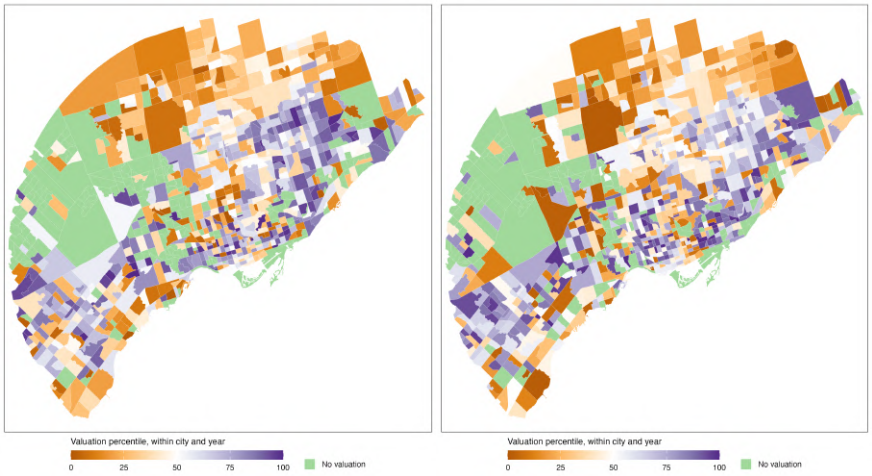
Notes: The figure plots the kernel density of immigrant enclaves as a function of distance from the Central Business District, pooling census tracts across Montréal, Toronto, and Vancouver. A census tract is classified as an enclave if immigrants comprise at least 50% of its resident population. The vertical lines represent the average distance from the CBD for each year. The 1991 immigrant share is projected onto 2021 census geographies using the Canadian Longitudinal Tract Database (Allen and Taylor, 2018). Distance is measured in kilometers from the centroid of each census tract to the Central Business District.

Figure 7: Neighborhood Valuations for Toronto in 2021



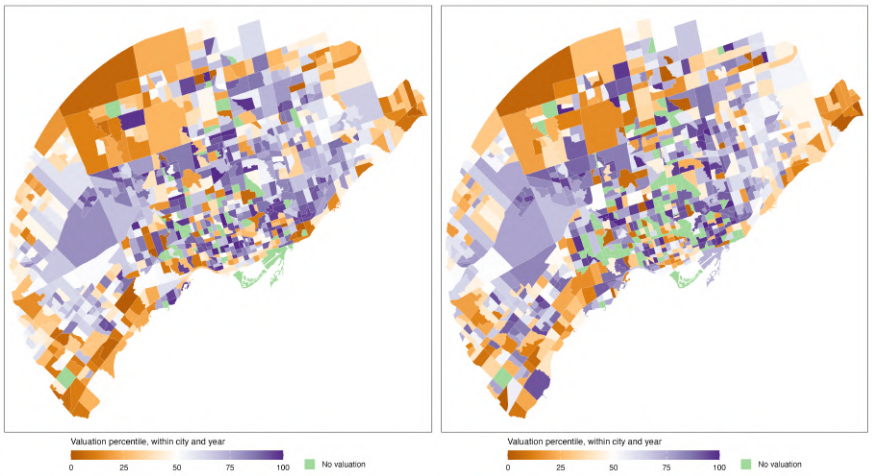
(a) Rich Natives

(b) Poor Natives



(c) Rich East Asians

(d) Poor East Asians

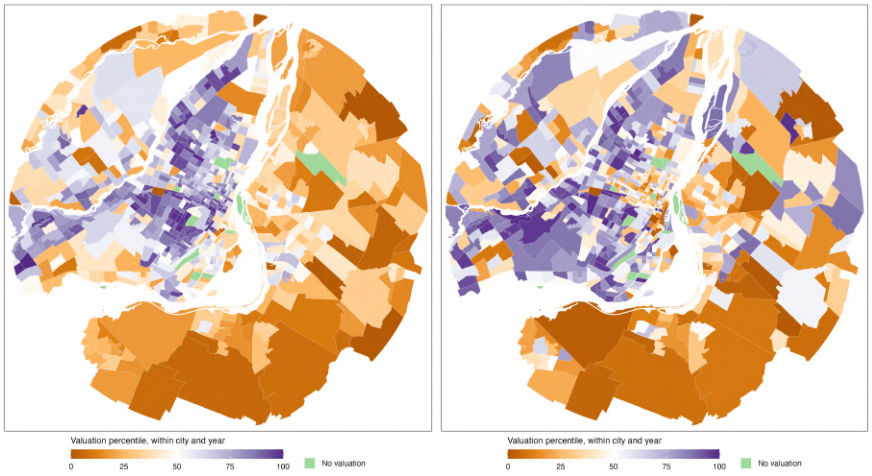


(e) Rich South Asians

(f) Poor South Asians

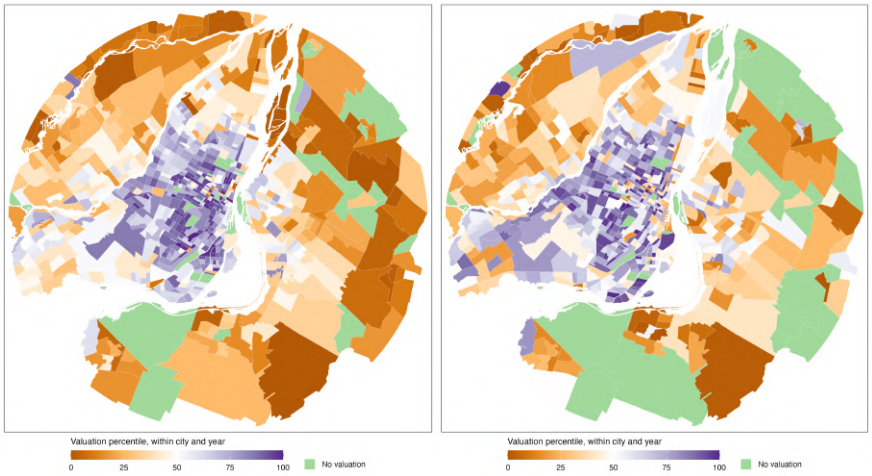
Notes: These maps plot the 2021 valuation percentile for natives and the two largest immigrant groups in Toronto, split by income. Tracts in green receive no valuation for the group (see Section 2.4). We cropped the metropolitan area to show census tracts within a radius from the city center that includes 75% of the MSA population.

Figure 8: Neighborhood Valuations for Montréal in 2021



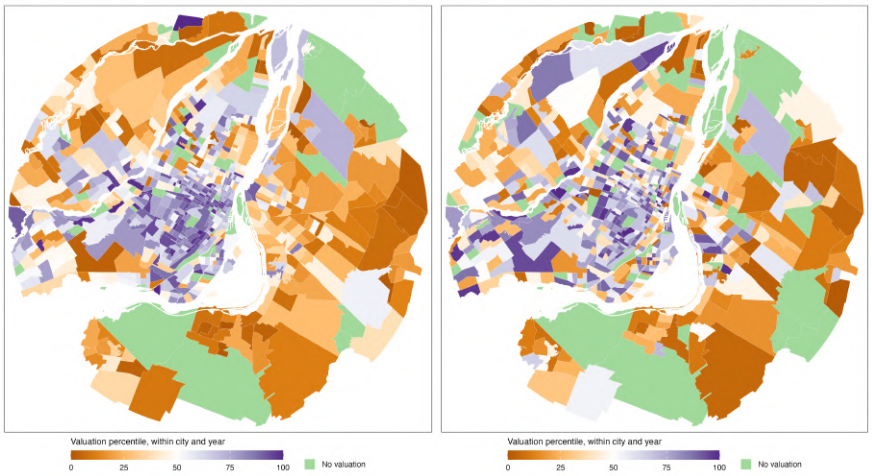
(a) Rich Natives

(b) Poor Natives



(c) Rich East Asians

(d) Poor East Asians

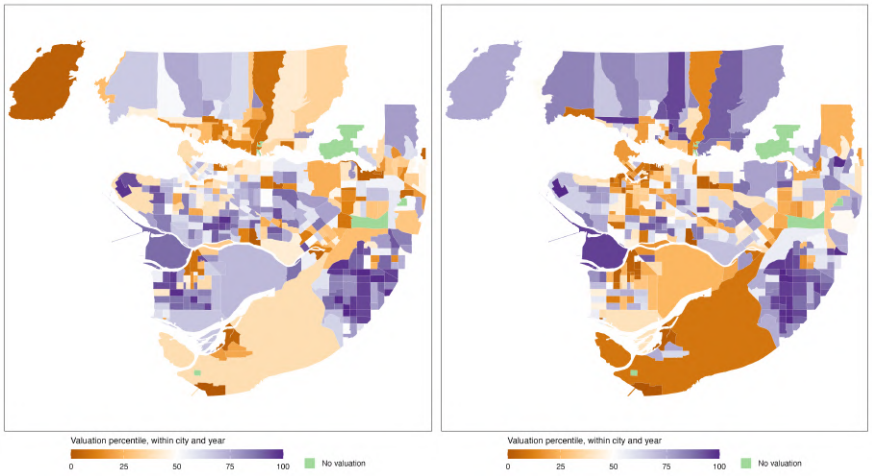


(e) Rich South Asians

(f) Poor South Asians

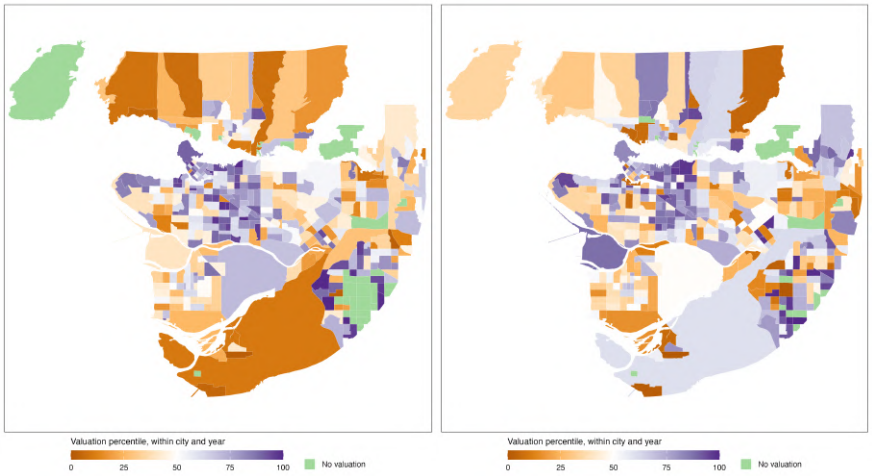
Notes: These maps plot the 2021 valuation percentile for natives and the two largest immigrant groups in Montréal, split by income. Tracts in green receive no valuation for the group (see Section 2.4). We cropped the metropolitan area to show census tracts within a radius from the city center that includes 75% of the MSA population.

Figure 9: Neighborhood Valuations for Vancouver in 2021



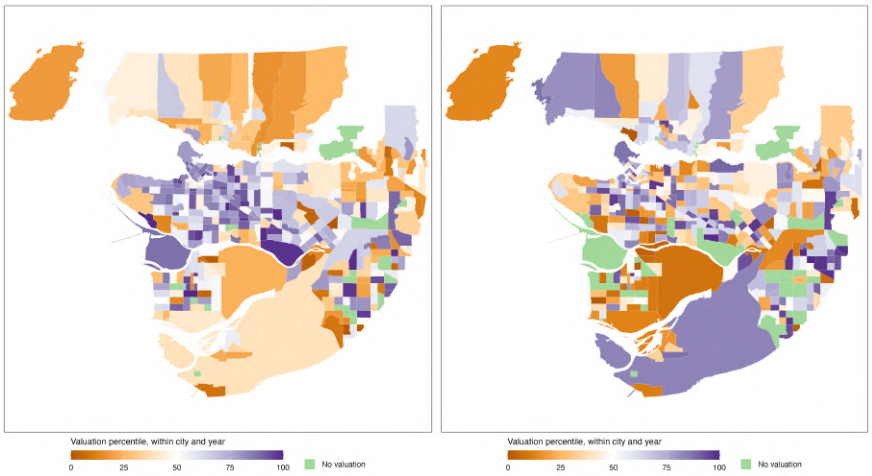
(a) Rich Natives

(b) Poor Natives



(c) Rich East Asians

(d) Poor East Asians



(e) Rich South Asians

(f) Poor South Asians

Notes: These maps plot the 2021 valuation percentile for natives and the two largest immigrant groups in Vancouver, split by income. Tracts in green receive no valuation for the group (see Section 2.4). We cropped the metropolitan area to show census tracts within a radius from the city center that includes 75% of the MSA population.

8 Tables

Table 1: Housing Expenditure Shares by City ($1 - \hat{\phi}_k$)

	Montréal		Toronto		Vancouver	
	Poor (1)	Rich (2)	Poor (3)	Rich (4)	Poor (5)	Rich (6)
Native-born	26.55%	12.69%	33.45%	16.46%	35.16%	18.13%
Europe & United States	29.65%	14.07%	33.69%	16.17%	34.82%	17.82%
Africa & Middle East	31.54%	14.90%	39.60%	19.28%	40.72%	21.25%
South Asia	31.71%	15.53%	39.58%	19.31%	41.18%	21.00%
East Asia	31.21%	14.74%	38.56%	17.87%	40.39%	20.00%
Oceania	32.51%	16.24%	32.74%	17.32%	37.06%	20.16%
Latin America	30.02%	14.69%	35.48%	17.79%	37.87%	19.24%

Note: Shares are computed from census microdata as the median ratio of shelter costs to household income, by group, city, and census year, then averaged across years. Households with zero shelter costs or zero income are excluded. Immigrants from Oceania are reported for completeness but are not part of the estimation sample.

Table 2: Correlates of Estimated Neighborhood Valuations – Rich

	Native	Western	Lat. Am.	Afr. & ME	E. Asia	S. Asia
Panel A: Correlation of $\log \hat{\hat{V}}^k$ *with rich native-born valuation*						
Correlation	1.000	0.385^a	0.413^a	0.487^a	0.304^a	0.417^a
		(0.020)	(0.019)	(0.018)	(0.024)	(0.020)
Panel B: Correlation of $\log \hat{\hat{V}}^k$ *with tract characteristics*						
Log rent	-0.106^a	-0.006	-0.096^a	-0.100^a	-0.151^a	-0.194^a
	(0.022)	(0.020)	(0.020)	(0.019)	(0.024)	(0.022)
Log median income	-0.266^a	-0.087^a	-0.175^a	-0.216^a	-0.188^a	-0.276^a
	(0.020)	(0.021)	(0.021)	(0.020)	(0.025)	(0.023)
Own-group share	-0.567^a	0.213^a	0.241^a	0.212^a	0.132^a	0.136^a
	(0.016)	(0.019)	(0.017)	(0.016)	(0.021)	(0.019)
Log distance to CBD	-0.249^a	-0.383^a	-0.358^a	-0.392^a	-0.260^a	-0.430^a
	(0.021)	(0.018)	(0.020)	(0.020)	(0.024)	(0.023)

Note: All values in the table correspond to 2021. The sample includes valuations for rich individuals, pooling the three cities together. All values are standardized to have a mean of zero and a standard deviation of one. The correlations correspond to the regression coefficient between valuation and the neighborhood characteristic, controlling for city fixed effects. Robust standard errors in parentheses. $^a \leq 0.01$, $^b \leq 0.05$, and $^c \leq 0.1$.

Table 3: Correlates of Estimated Neighborhood Valuations – Poor

	Native	Western	Lat. Am.	Afr. & ME	E. Asia	S. Asia
Panel A: Correlation of $\log \hat{\hat{V}}^k$ *with poor native-born valuation*						
Correlation	1.000	0.211^a	0.292^a	0.293^a	0.226^a	0.262^a
		(0.021)	(0.020)	(0.020)	(0.023)	(0.022)
Panel B: Correlation of $\log \hat{\hat{V}}^k$ *with tract characteristics*						
Log rent	-0.206^a	-0.014	-0.141^a	-0.072^a	-0.150^a	-0.174^a
	(0.023)	(0.021)	(0.021)	(0.020)	(0.021)	(0.024)
Log median income	-0.182^a	-0.067^a	-0.168^a	-0.145^a	-0.155^a	-0.228^a
	(0.022)	(0.022)	(0.021)	(0.019)	(0.024)	(0.023)
Own-group share	-0.301^a	0.169^a	0.233^a	0.184^a	0.080^a	0.135^a
	(0.020)	(0.020)	(0.017)	(0.016)	(0.019)	(0.018)
Log distance to CBD	-0.026	-0.215^a	-0.202^a	-0.335^a	-0.239^a	-0.328^a
	(0.024)	(0.020)	(0.024)	(0.022)	(0.023)	(0.028)

Note: All values presented in the table correspond to 2021. The sample includes valuations for poor individuals, pooling the three cities together. All values are standardized to have a mean of zero and a standard deviation of one. The correlations correspond to the regression coefficient between valuation and the neighborhood characteristic, controlling for city fixed effects. Robust standard errors in parentheses. $^a \leq 0.01$, $^b \leq 0.05$, and $^c \leq 0.1$.

Table 4: Responses to Neighborhood Composition

	Responding group k											
	Native		Western		Lat. Am.		Afr. & ME		E. Asia		S. Asia	
	Poor (1)	Rich (2)	Poor (3)	Rich (4)	Poor (5)	Rich (6)	Poor (7)	Rich (8)	Poor (9)	Rich (10)	Poor (11)	Rich (12)
Response to share of:												
Western countries	-0.445^{a}	-0.326^{a}	0.786^{a}	0.804^{a}	-0.284^{b}	-0.108	-0.067	0.080	-0.078	0.018	0.190	0.062
	(0.111)	(0.112)	(0.112)	(0.100)	(0.124)	(0.110)	(0.121)	(0.090)	(0.114)	(0.109)	(0.120)	(0.109)
Latin America & Caribbean	0.183	0.268^{b}	0.187	-0.108	0.590^{a}	0.338^{a}	0.483^{a}	0.178	0.362^{c}	0.479^{b}	0.208	-0.186
	(0.123)	(0.115)	(0.173)	(0.170)	(0.142)	(0.129)	(0.121)	(0.126)	(0.189)	(0.189)	(0.136)	(0.140)
Africa & Middle East	-0.073	0.406^{a}	0.028	0.290^{a}	-0.135	-0.013	0.328^{a}	0.674^{a}	-0.254^{b}	0.031	0.101	0.081
	(0.106)	(0.086)	(0.106)	(0.106)	(0.120)	(0.105)	(0.092)	(0.088)	(0.105)	(0.106)	(0.108)	(0.099)
East Asia	-0.125^{c}	0.258^{a}	0.052	0.045	-0.116	-0.170^{b}	0.017	0.108^{b}	0.104^{c}	0.264^{a}	0.166^{a}	0.026
	(0.065)	(0.051)	(0.072)	(0.068)	(0.075)	(0.068)	(0.058)	(0.055)	(0.057)	(0.058)	(0.063)	(0.057)
Southern Asia	0.225^{a}	0.355^{a}	0.099	0.018	-0.195^{b}	-0.144^{b}	-0.001	0.009	-0.030	0.123	0.228^{a}	0.269^{a}
	(0.065)	(0.063)	(0.087)	(0.083)	(0.079)	(0.072)	(0.071)	(0.066)	(0.086)	(0.085)	(0.060)	(0.058)
Kleibergen–Paap F	239.78	239.78	239.78	239.78	239.78	239.78	239.78	239.78	239.78	239.78	239.78	239.78
Observations	3,145	3,145	3,096	3,084	3,075	3,080	3,137	3,106	3,054	2,882	3,036	2,996

Note: Each column reports 2SLS estimates of $\mathbf{\Gamma}_b^k$ from equation (18). The dependent variable is the price-adjusted log PageRank valuation. The native-born share is the omitted category. Instruments are tract shares two periods (ten years) earlier; controls are all groups' price-adjusted valuations one period (five years) earlier, the share of rich households, and group-city and period fixed effects. Standard errors clustered at the census-tract level in parentheses. $^{a} \leq 0.01$, $^{b} \leq 0.05$, and $^{c} \leq 0.1$.

References

Ahlfeldt, Gabriel M, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf, "The economics of density: Evidence from the Berlin Wall," *Econometrica*, 2015, *83* (6), 2127–2189.

Albouy, David, "What are cities worth? Land rents, local productivity, and the total value of amenities," *Review of Economics and Statistics*, 2016, *98* (3), 477–487.

Allen, Jeff and Zack Taylor, "A new tool for neighbourhood change research: The Canadian Longitudinal Census Tract Database, 1971–2016," *The Canadian Geographer/Le Géographe canadien*, 2018.

Almagro, Milena and Tomás Domínguez-Iino, "Location Sorting and Endogenous Amenities: Evidence From Amsterdam," *Econometrica*, 2025, *93* (3), 1031–1071.

Altonji, Joseph G. and David Card, "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives," in "Immigration, Trade, and the Labor Market," University of Chicago Press, 1991, pp. 201–234.

Arcidiacono, Peter and Robert A. Miller, "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity," *Econometrica*, 2011, *79* (6), 1823–1867.

Bagagli, Sara, "The (express) way to segregation: Evidence from Chicago," 2025.

Bartel, Ann P, "Where do the new US immigrants live?," *Journal of labor economics*, 1989, *7* (4), 371–391.

Bayer, Patrick and Christopher Timmins, "On the equilibrium properties of locational sorting models," *Journal of Urban Economics*, 2005, *57* (3), 462–477.

_ , Fernando Ferreira, and Robert McMillan, "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," *Journal of Political Economy*, 2007, *115* (4), 588–638.

_ , Robert McMillan, Alvin Murphy, and Christopher Timmins, "A Dynamic Model of Demand for Houses and Neighborhoods," *Econometrica*, 2016, *84* (3), 893–942.

_ , _ , and Kim Rueben, "An Equilibrium Model of Sorting in an Urban Housing Market," Working Paper 10865, National Bureau of Economic Research 2004.

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

Beauregard, Pierre-Loup, Alain-G Gagnon, and Jean-Denis Garon, "Managing immigration in the Canadian federation: The case of Quebec," in "International affairs and Canadian migration policy," Springer, 2020, pp. 227–245.

Bell, Alex, Sophie Calder-Wang, and Shusheng Zhong, "Measuring housing quality using revealed preference: A geographic pagerank approach," Technical Report, National Bureau of Economic Research 2026.

_ , Stephen B Billings, Sophie Calder-Wang, and Shusheng Zhong, "An anti-iv approach for pricing residential amenities: Applications to flood risk," *The Wharton School Research Paper*, 2025.

Caetano, Gregorio and Vikram Maheshri, "A unified empirical framework to study neighborhood segregation," *Quantitative Economics*, 2025, *16* (3), 1023–1057.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon, "The Production Function for Housing: Evidence from France," *Journal of Political Economy*, 2021, *129* (10), 2766–2816.

Couture, Victor, Cecile Gaubert, Jessie Handbury, and Erik Hurst, "Income Growth and the Distributional Effects of Urban Spatial Sorting," *The Review of Economic Studies*, 2023. rdad048.

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

Davis, Donald R, Matthew Easton, and Stephan Thies, "Segregation, Spillovers, and the Locus of Racial Change," Working Paper 34246, National Bureau of Economic Research 2025.

Dawkins, Casey J, "Exploring recent trends in immigrant suburbanization," *Cityscape*, 2009, pp. 81–97.

Diamond, Rebecca, "The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000," *American Economic Review*, 2016, *106* (3), 479–524.

Dustmann, Christian, Tommaso Frattini, and Ian P Preston, "The effect of immigration along the distribution of wages," *Review of Economic Studies*, 2013, *80* (1), 145–173.

Edin, Per-Anders, Peter Fredriksson, and Olof Åslund, "Ethnic enclaves and the economic success of immigrants—Evidence from a natural experiment," *The Quarterly Journal of Economics*, 2003, *118* (1), 329–357.

Epple, Dennis, Brett Gordon, and Holger Sieg, "A New Approach to Estimating the Production Function for Housing," *American Economic Review*, 2010, *100* (3), 905–24.

Ferreira, Fernando V and Maisy Wong, "Estimating Preferences for Neighborhood Amenities Under Imperfect Information," Working Paper 28165, National Bureau of Economic Research 2020.

Fogel, Jamie, "Valuing American Cities: A Revealed Preference Approach," 2021.

Logan, John R, Richard D Alba, and Wenquan Zhang, "Immigrant enclaves and ethnic communities in New York and Los Angeles," *American sociological review*, 2002, *67* (2), 299–322.

Massey, Douglas S, "American apartheid: Segregation and the making of the underclass," *American journal of sociology*, 1990, *96* (2), 329–357.

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

Newstead, Nick, Kaitlyn Hobbs, Cal Giunta, and Sheldon Birkett, "Active and Public Transportation Spatial Accessibility Measures: Methodology and Key Results," Reports on Special Business Projects 18-001-X2024005, Statistics Canada, Ottawa, Ontario November 2024.

Page, Lawrence, Sergey Brin, Rajeev Motwani, and Terry Winograd, "The PageRank citation ranking: Bringing order to the web.," Technical Report, Stanford infolab 1999.

Saiz, Albert and Susan Wachter, "Immigration and the Neighborhood," *American Economic Journal: Economic Policy*, May 2011, *3* (2), 169–88.

Sorkin, Isaac, "Ranking Firms Using Revealed Preference," *The Quarterly Journal of Economics*, 2018, *133* (3), 1331–1393.

Statistics Canada, "Census Profile. 2021 Census of Population," Statistics Canada Catalogue number 98-316-X2021001 2023. Released November 15, 2023. Accessed October 11, 2026.

_ , "Spatial Access Measures," Statistics Canada Catalogue no. 27-26-0001 July 2023. Developed in collaboration with Infrastructure Canada. Estimates for access to places of employment corrected August 22, 2023.

United Nations, "International Migrant Stock 2024: Key Facts and Figures," Technical Report UN DESA/POP/2024/DC/NO. 13, United Nations 2024.

A Model Derivations and Equilibrium

A.1 Derivation of the Indirect Utility

The representative type-k household in neighborhood j solves

$$\max_{c,\,h}\; b_{jt}^{k}\left(\frac{c}{\phi_k}\right)^{\phi_k}\left(\frac{h}{1-\phi_k}\right)^{1-\phi_k} \qquad \text{s.t.} \quad c + Q_{jt}\,h = w_{kt}.$$

The Lagrangian is $\mathcal{L} = b_{jt}^{k}\,(\phi_k)^{-\phi_k}(1-\phi_k)^{-(1-\phi_k)}\,c^{\phi_k}\,h^{1-\phi_k} - \lambda(c + Q_{jt}\,h - w_{kt})$. First-order conditions:

$$\frac{\partial \mathcal{L}}{\partial c} = b_{jt}^{k}\,(\phi_k)^{-\phi_k}(1-\phi_k)^{-(1-\phi_k)}\,\phi_k\,c^{\phi_k-1}\,h^{1-\phi_k} = \lambda, \tag{A.1}$$

$$\frac{\partial \mathcal{L}}{\partial h} = b_{jt}^{k}\,(\phi_k)^{-\phi_k}(1-\phi_k)^{-(1-\phi_k)}\,(1-\phi_k)\,c^{\phi_k}\,h^{-\phi_k} = \lambda\,Q_{jt}. \tag{A.2}$$

Dividing (A.2) by (A.1):

$$\frac{(1-\phi_k)\,c}{\phi_k\,h} = Q_{jt} \quad\Longrightarrow\quad c = \frac{\phi_k}{1-\phi_k}\,Q_{jt}\,h.$$

Substituting into the budget constraint:

$$\frac{\phi_k}{1-\phi_k}\,Q_{jt}\,h + Q_{jt}\,h = w_{kt} \quad\Longrightarrow\quad h^{*} = \frac{(1-\phi_k)\,w_{kt}}{Q_{jt}}, \qquad c^{*} = \phi_k\,w_{kt}.$$

Substituting (c^{*}, h^{*}) into the objective:

$$\begin{aligned}
\bar{u}_t^{k}(j) &= b_{jt}^{k}\left(\frac{\phi_k\,w_{kt}}{\phi_k}\right)^{\phi_k}\left(\frac{(1-\phi_k)\,w_{kt}}{(1-\phi_k)\,Q_{jt}}\right)^{1-\phi_k} \\
&= b_{jt}^{k}\,w_{kt}^{\phi_k}\,w_{kt}^{1-\phi_k}\,Q_{jt}^{-(1-\phi_k)} \\
&= b_{jt}^{k}\,w_{kt}\,Q_{jt}^{\phi_k-1}.
\end{aligned} \tag{A.3}$$

This is equation (5) of the main text.

A.2 Housing Supply

To close the model, floor space in each neighborhood is produced by a competitive construction sector using fixed geographic land T_j and mobile capital L_{jt} as inputs (Combes et al., 2021; Epple et al., 2010). The

Cobb-Douglas production function is

$$H_{jt} \;=\; T_j^{\mu}\, L_{jt}^{1-\mu}, \tag{A.4}$$

where $\mu \in (0,1)$ is the land share.[5] Capital trades at a common price p and land earns rent r_{jt}.

Cost minimization. The cost of producing H_{jt} units is

$$C(H_{jt}) = \min_{L_{jt}} \; r_{jt}\, T_j + p\, L_{jt} \quad \text{s.t.} \quad T_j^{\mu}\, L_{jt}^{1-\mu} = H_{jt}.$$

Since T_j is fixed, the firm chooses L_{jt} to meet the output target:

$$L_{jt} = \left(\frac{H_{jt}}{T_j^{\mu}} \right)^{1/(1-\mu)}. \tag{A.5}$$

Total cost is therefore

$$C(H_{jt}) = r_{jt}\, T_j + p \left(\frac{H_{jt}}{T_j^{\mu}} \right)^{1/(1-\mu)}, \tag{A.6}$$

and marginal cost is

$$\frac{\partial C}{\partial H_{jt}} = \frac{p}{1-\mu}\, T_j^{-\mu/(1-\mu)}\, H_{jt}^{\mu/(1-\mu)}. \tag{A.7}$$

Competitive pricing. Setting the price of floor space equal to marginal cost, $Q_{jt} = \partial C / \partial H_{jt}$, and solving for H_{jt}:

$$\begin{aligned} Q_{jt} &= \frac{p}{1-\mu}\, T_j^{-\mu/(1-\mu)}\, H_{jt}^{\mu/(1-\mu)} \\ \Longrightarrow \quad H_{jt}^{\mu/(1-\mu)} &= \frac{(1-\mu)\, Q_{jt}}{p}\, T_j^{\mu/(1-\mu)} \\ \Longrightarrow \quad H_{jt} &= \left(\frac{1-\mu}{p} \right)^{(1-\mu)/\mu} T_j\, Q_{jt}^{(1-\mu)/\mu}. \end{aligned}$$

Defining $\varphi(\mu) \equiv \big((1-\mu)/p\big)^{(1-\mu)/\mu}$, the housing supply function is

$$H_{jt}^{S} \;=\; \varphi(\mu)\, T_j\, Q_{jt}^{(1-\mu)/\mu}, \tag{A.8}$$

with implied housing supply elasticity $\partial \log H_{jt}^{S} / \partial \log Q_{jt} = (1-\mu)/\mu$.

[5] Geographic land in each neighborhood is fixed, consistent with the use of constant census-tract boundaries.

Housing demand. A type-k household in neighborhood j demands $h_{jt}^k = (1-\phi_k)\, w_{kt}/Q_{jt}$ units of floor space (Appendix A.1), so aggregate housing demand is

$$H_{jt}^D \;=\; \sum_k M_{jt}^k \,\frac{(1-\phi_k)\, w_{kt}}{Q_{jt}}. \tag{A.9}$$

A.3 Stationary Equilibrium

Population composition, housing prices, and amenities are simultaneously determined in equilibrium.

Law of motion. Let $\Pi^k(Q_t,\, b_t)$ be the $(J+1)\times(J+1)$ transition matrix whose entry in row j and column h is the choice probability $P_t^k(j \mid h)$ of equation (10). Stacking the type-k population across locations in the column vector $M_t^k = [M_{0t}^k,\, M_{1t}^k, \ldots, M_{Jt}^k]'$, where M_{0t}^k is the mass of type-k households in the outside option, the population evolves according to

$$M_t^k \;=\; \Pi^k(Q_t,\, b_t)\, M_{t-1}^k. \tag{A.10}$$

Since households only move between neighborhoods and the outside option, the total mass $\bar{M}^k = \sum_{j=0}^J M_{jt}^k$ of each type is constant.

Definition 1 (Stationary Population Distribution)**.** *Given prices* $\mathbf{Q} = (Q_1, \ldots, Q_J)'$ *and amenities* $\mathbf{b} = \{b_j^k\}_{k,j}$*, a* stationary population distribution $M(\mathbf{Q}, \mathbf{b}) = \{M^k(\mathbf{Q}, \mathbf{b})\}_k$ *satisfies, for every type* k*,*

$$M^k(\mathbf{Q}, \mathbf{b}) \;=\; \Pi^k(\mathbf{Q}, \mathbf{b})\, M^k(\mathbf{Q}, \mathbf{b}), \qquad \sum_{j=0}^{J} M_j^k(\mathbf{Q}, \mathbf{b}) = \bar{M}^k, \tag{A.11}$$

where the choice probabilities in Π^k *are computed with the continuation values* EV^k *of the stationary dynamic program at* $(\mathbf{Q}, \mathbf{b})$*.*

Definition 2 (Stationary Equilibrium)**.** *A* stationary equilibrium *consists of (i) housing prices* $\mathbf{Q}$*, (ii) amenities* $\mathbf{b}$*, and (iii) a stationary population distribution* $M(\mathbf{Q}, \mathbf{b})$ *satisfying* (A.11)*, such that:*

1. ***Floor-space market clearing:*** *for all* $j \in \{1, \ldots, J\}$*, housing demand* (A.9) *equals housing sup-*

ply (A.8),

$$\sum_k M_j^k \frac{(1-\phi_k)\, w_k}{Q_j} \;=\; \varphi(\mu)\, T_j\, Q_j^{(1-\mu)/\mu}.$$

2. ***Amenity consistency:*** *for all* $j \in \{1, \ldots, J\}$ *and all* k,

$$b_j^k \;=\; G^k(M_j),$$

where $M_j = [M_j^1, \ldots, M_j^K, M_j^N]'$ *is the composition of neighborhood* j*, as in equation* (1).

Existence. A stationary equilibrium exists under a finite capacity constraint $H_j \leq \bar{H} < \infty$ for all j, which prevents unbounded sorting into a single location. The argument follows Almagro and Domínguez-Iino (2025). Consider the mapping $\Phi\colon (\mathbf{Q}, \mathbf{b}) \mapsto (\mathbf{Q}', \mathbf{b}')$ that (i) computes flow utilities from equation (6) and the stationary population distribution $M(\mathbf{Q}, \mathbf{b})$, which exists and is unique by Perron–Frobenius since the logit transition matrix has strictly positive entries; (ii) sets $b_j'^k = G^k(M_j)$; and (iii) sets Q_j' to clear the floor-space market in each j. The capacity constraint bounds prices from above, positive marginal cost bounds them from below, and the boundedness of G^k bounds amenities, so Φ maps a compact, convex set into itself. Each step of Φ is continuous: the logit probabilities are smooth in $(\mathbf{Q}, \mathbf{b})$, the stationary distribution is a continuous function of the transition matrix, G^k is continuous by assumption, and the market-clearing price is a continuous function of M because excess demand is strictly decreasing in Q_j. By Brouwer's fixed-point theorem, Φ has a fixed point, which is a stationary equilibrium.

Uniqueness. Uniqueness is *not* guaranteed when amenities are endogenous. The feedback $b_j^k = G^k(M_j)$ can generate multiple fixed points. In a segregated equilibrium, immigrant group k concentrates in a few neighborhoods, which raises its amenities there and reinforces the concentration. In an integrated equilibrium, groups are dispersed, amenities are moderate everywhere, and no group has a strong incentive to cluster. Both configurations can satisfy all equilibrium conditions, consistent with the observation that similar cities exhibit different degrees of residential segregation. The estimation in Section 5 does not require selecting among equilibria: it conditions on the observed flows and prices, which are generated by whichever equilibrium is realized in the data.

B Estimation Derivations

B.1 Identification of Neighborhood Valuations via PageRank

This section derives formally that the type-specific PageRank algorithm applied to within-city bilateral migration flows recovers the neighborhood valuation $\widetilde{V}^k(j)$ up to a scalar normalization, with an approximation error of order $O(J^{-1})$.

B.1.1 Setup and Notation

Let $\mathcal{J} = \{1, \ldots, J\}$ be the set of within-city neighborhoods and $j = 0$ the outside option. Fix a household type k and a time period t; we suppress t throughout.

Define the exponentiated neighborhood value:

$$V^k(j) \equiv e^{u^k(j) + \beta\, EV^k(j)}, \qquad j \in \mathcal{J}, \tag{B.1}$$

and the aggregates

$$A^k \equiv \sum_{j' \in \mathcal{J}} V^k(j'), \tag{B.2}$$

$$X^k \equiv e^{u^k(0) - m_2^k + \beta\, EV^k(0)}. \tag{B.3}$$

The target object is the flow-relevant value $\widetilde{V}^k(j) \equiv V^k(j)/M_j^k$.

B.1.2 Within-City Bilateral Flows

PageRank is applied exclusively to within-city movers. Within $\mathcal{J}$, moving costs are symmetric: $MC^k(j,h) = MC^k(h,j) = m_1^k$ for all $j \neq h$. The expected flow from h to j ($j \neq h$) is

$$\Upsilon_{h \to j}^k = M_h^k \cdot \frac{e^{-m_1^k}\, V^k(j)}{D^k(h)}, \tag{B.4}$$

where $D^k(h)$ is the full logit denominator at origin h.

B.1.3 Expanding the Denominator

Separating stay, move, and exit terms:

$$\begin{aligned} D^k(h) &= V^k(h) + e^{-m_1^k}\big(A^k - V^k(h)\big) + X^k \\ &= \underbrace{e^{-m_1^k}A^k + X^k}_{\text{common}} + \underbrace{\big(1 - e^{-m_1^k}\big)\,V^k(h)}_{\text{origin-specific stay bonus}}. \end{aligned} \tag{B.5}$$

B.1.4 The Denominator Ratio

We show that the denominator ratio is $1 + O(J^{-1})$.

Lemma 1 (Denominator ratio)**.** *Suppose $v_{\min} \le V^k(j') \le v_{\max}$ for all $j' \in \mathcal{J}$, with $v_{\min} > 0$ and $v_{\max} < \infty$ independent of J. Then for any $j, h \in \mathcal{J}$:*

$$\frac{D^k(j)}{D^k(h)} = 1 + O(J^{-1}), \tag{B.6}$$

with implicit constant $C(m_1^k) = 2(1 - e^{-m_1^k})v_{\max}/(e^{-m_1^k}v_{\min})$.

Proof. From (B.5): $D^k(j) - D^k(h) = (1 - e^{-m_1^k})[V^k(j) - V^k(h)]$. Dividing by $D^k(h) \ge e^{-m_1^k}A^k \ge e^{-m_1^k}J\,v_{\min}$:

$$\left|\frac{D^k(j)}{D^k(h)} - 1\right| \;\le\; \frac{(1 - e^{-m_1^k}) \cdot 2v_{\max}}{e^{-m_1^k} \cdot J \cdot v_{\min}} \;=\; \frac{C(m_1^k)}{J}. \tag{B.7}$$

□

Intuition. When J is large, the common term $e^{-m_1^k}A^k = O(J)$ dominates the origin-specific stay bonus $O(1)$, making the denominator approximately equal across origins.

B.1.5 The Bilateral Flow Ratio

Dividing $\Upsilon^k_{h\to j}$ by $\Upsilon^k_{j\to h}$:

$$\frac{\Upsilon^k_{h\to j}}{\Upsilon^k_{j\to h}} = \frac{M^k_h}{M^k_j}\cdot\frac{V^k(j)}{V^k(h)}\cdot\underbrace{\frac{e^{-m^k_1}}{e^{-m^k_1}}}_{=1}\cdot\underbrace{\frac{D^k(j)}{D^k(h)}}_{=1+O(J^{-1})}$$

$$= \frac{\widetilde{V}^k(j)}{\widetilde{V}^k(h)}\cdot\big[1+O(J^{-1})\big]. \tag{B.8}$$

B.1.6 Cross-Multiplication and the PageRank Fixed Point

Ignoring the $O(J^{-1})$ remainder: $\Upsilon^k_{h\to j}\,\widetilde{V}^k(h) \approx \Upsilon^k_{j\to h}\,\widetilde{V}^k(j)$. Summing over $h \in \mathcal{J}$ and dividing by total out-flow:

$$\widetilde{V}^k(j) \;=\; \frac{\displaystyle\sum_{h\in\mathcal{J}} \Upsilon^k_{h\to j}\,\widetilde{V}^k(h)}{\displaystyle\sum_{h\in\mathcal{J}} \Upsilon^k_{j\to h}}, \qquad \forall\, j\in\mathcal{J}. \tag{B.9}$$

This is the PageRank fixed-point equation.

B.1.7 Matrix Form, Existence, and Identification

Let $\boldsymbol{\Upsilon}^k$ be the $J\times J$ flow matrix whose entry in row j and column h is $\Upsilon^k_{h\to j}$ (destinations index rows), and let $\mathbf{S}^k = \mathrm{diag}(S^k_1,\ldots,S^k_J)$ with $S^k_j = \sum_h \Upsilon^k_{j\to h}$ the total out-flow from j. Equation (B.9) in matrix form:

$$\widetilde{\mathbf{V}}^k \;=\; (\mathbf{S}^k)^{-1}\,\boldsymbol{\Upsilon}^k\,\widetilde{\mathbf{V}}^k. \tag{B.10}$$

Each column of $\boldsymbol{\Upsilon}^k$ sums to the total out-flow of the corresponding origin, so $\boldsymbol{\Upsilon}^k(\mathbf{S}^k)^{-1}$ is column-stochastic: it is the transition matrix of the Markov chain of within-city movers, and it has an eigenvalue equal to one, associated with the stationary distribution $\boldsymbol{\pi}^k$ of that chain. Since $(\mathbf{S}^k)^{-1}\boldsymbol{\Upsilon}^k = (\mathbf{S}^k)^{-1}\big[\boldsymbol{\Upsilon}^k(\mathbf{S}^k)^{-1}\big]\mathbf{S}^k$, the two matrices share their eigenvalues, and the solution of (B.10) is

$$\widetilde{\mathbf{V}}^k \;\propto\; (\mathbf{S}^k)^{-1}\,\boldsymbol{\pi}^k. \tag{B.11}$$

In words, the valuation of a neighborhood is its weight in the stationary distribution of movers divided by its total out-flow. By Perron–Frobenius, existence and uniqueness (up to scalar) hold if the neighborhood

graph is strongly connected. We normalize $\sum_j \widetilde{V}^k(j) = J$.

The total approximation error in $\widetilde{\mathbf{V}}^k$ is $O(J^{-1})$ uniformly in j. The implicit constant $C(m_1^k)$ in Lemma 1 increases with the within-city moving cost and with the dispersion of valuations, $v_{\max}/v_{\min}$, so the approximation is most accurate when moving costs are moderate relative to the number of neighborhoods.

B.2 Estimation of Moving Costs

The moving costs m_1^k and m_2^k are not needed for the results in the main text, but they are needed to solve and simulate the model. This section shows that they can be recovered from the same flows used in the PageRank step, combining the PageRank valuations with ratios of choice probabilities, in which continuation values cancel (Arcidiacono and Miller, 2011). Throughout, hats denote sample analogues: $\hat{P}^k(j|h)$ is the observed share of type-k households at h at the start of a period who are at j at the end of it, and $\hat{\widetilde{V}}^k(j)$ is the PageRank valuation of Section 5.2.

Within-city moving cost. Consider a household of type k at origin $h \in \mathcal{J}$ choosing between moving to $j \in \mathcal{J}$, $j \neq h$, at cost m_1^k and staying at h at no cost. By equation (10), the logit denominator cancels in the ratio of the two choice probabilities:

$$\frac{P^k(j|h)}{P^k(h|h)} = \frac{\exp\{u^k(j) - m_1^k + \beta\, EV^k(j)\}}{\exp\{u^k(h) + \beta\, EV^k(h)\}} = \frac{\widetilde{V}^k(j)\, M_j^k}{\widetilde{V}^k(h)\, M_h^k}\, e^{-m_1^k}, \tag{B.12}$$

where the second equality uses $e^{u^k(j)+\beta EV^k(j)} = \widetilde{V}^k(j)\, M_j^k$ from equation (12). Summing (B.12) over within-city destinations and weighting by the origin population gives a condition on expected counts of movers and stayers:

$$\sum_h M_h^k \Big[\sum_{j \in \mathcal{J},\, j \neq h} \hat{P}^k(j|h) \;-\; \hat{P}^k(h|h)\, R_h^k\, e^{-m_1^k} \Big] = 0, \qquad R_h^k \equiv \frac{\sum_{j \in \mathcal{J},\, j \neq h} \hat{\widetilde{V}}^k(j)\, M_j^k}{\hat{\widetilde{V}}^k(h)\, M_h^k}, \tag{B.13}$$

where M_h^k is the origin population at the start of the period, and the sum also runs over periods. We estimate (B.13) by Poisson pseudo-maximum likelihood, which has the closed form

$$\hat{m}_1^k \;=\; \log \frac{\sum_h M_h^k\, \hat{P}^k(h|h)\, R_h^k}{\sum_h M_h^k \sum_{j \in \mathcal{J},\, j \neq h} \hat{P}^k(j|h)}. \tag{B.14}$$

Working with expected counts rather than log ratios keeps origins with no observed moves and avoids selecting on the outcome. Replacing $R_h^k e^{-m_1^k}$ with $\exp\{\gamma_0 + \gamma_1 \log R_h^k\}$ yields a specification test, since the model implies $\gamma_1 = 1$. The test checks the one-for-one response of out-flows to relative valuations; because R_h^k aggregates over destinations, it is not informative about moving costs that vary across origin–destination pairs.

Exit and entry costs. Define the flow-relevant value of the outside option analogously, $\widetilde{V}^k(0) \equiv e^{u^k(0)+\beta EV^k(0)}/M_0^k$. The same steps applied to exits from h (at cost m_2^k) and to entries from the outside option into h (at cost m_{entry}^k) give

$$\sum_h M_h^k \Big[\hat{P}^k(0|h) - \hat{P}^k(h|h)\, \frac{\widetilde{V}^k(0)\, M_0^k}{\hat{\widetilde{V}}^k(h)\, M_h^k}\, e^{-m_2^k}\Big] = 0, \tag{B.15}$$

$$\sum_h M_0^k \Big[\hat{P}^k(h|0) - \hat{P}^k(0|0)\, \frac{\hat{\widetilde{V}}^k(h)\, M_h^k}{\widetilde{V}^k(0)\, M_0^k}\, e^{-m_{\text{entry}}^k}\Big] = 0, \tag{B.16}$$

where $\hat{P}^k(0|0) = 1 - \sum_h \hat{P}^k(h|0)$, since the outside option is defined as the residual. The entry condition pins down $\widetilde{V}^k(0)M_0^k$ (period by period, up to $e^{m_{\text{entry}}^k}$). Substituting it into the exit condition eliminates both the outside value and the PageRank normalization, and identifies the round-trip cost. For a single period,

$$\hat{m}_2^k + \hat{m}_{\text{entry}}^k \;=\; \log \frac{\Big[\sum_h \dfrac{\hat{P}^k(h|h)}{\hat{\widetilde{V}}^k(h)}\Big]\, \hat{P}^k(0|0)\, \sum_h \hat{\widetilde{V}}^k(h) M_h^k}{\Big[\sum_h M_h^k \hat{P}^k(0|h)\Big]\, \sum_h \hat{P}^k(h|0)}. \tag{B.17}$$

With several periods, we allow $\widetilde{V}^k(0)$ to vary by period and estimate a common round-trip cost by Poisson pseudo-maximum likelihood.

The data alone do not separate the round trip into its two legs: the transformation $\big(\widetilde{V}^k(0),\, m_2^k,\, m_{\text{entry}}^k\big) \mapsto \big(\lambda\widetilde{V}^k(0),\, m_2^k + \log\lambda,\, m_{\text{entry}}^k - \log\lambda\big)$ leaves every choice probability unchanged. The model resolves this: equation (7) charges m_1^k for any move into a city neighborhood, including entry from the outside option, so $m_{\text{entry}}^k = m_1^k$ and $\hat{m}_2^k$ is the round-trip cost minus $\hat{m}_1^k$. Counterfactuals that hold $u^k(0)$ and moving costs fixed depend only on the round trip.

C Hedonic Valuation

In this section, we compare our neighborhood valuation estimated using flows between neighborhoods to the neighborhood valuation estimated from home sales in the city of Vancouver.

C.1 The British Columbia Assessment Data

We use the universe of residential property sales recorded in the province of British Columbia (BC) by the province's assessment agency, BCA. These records contain, in addition to home characteristics and sale price, the latitude and longitude of each property, which we use to identify the corresponding neighborhood. We use the 2021 census tract definition as our neighborhood definition. The data cover the period from 1979 to 2021 for the Vancouver metropolitan area. We convert each dollar amount into 2021 Canadian dollars. The final dataset contains 1,452,733 property sales that meet our criteria.[6]

Since we do not observe buyers' characteristics in the sales data, we aggregate neighborhood valuations and estimate one general valuation across all groups.

C.2 Empirical Strategy

The goal is to measure how much the market pays for a neighborhood, separately from what it pays for the dwelling. To do so, we estimate a hedonic price regression of property price on characteristics and neighborhood dummies.

Let i index a property, t a sale year, $b(t)$ the five-year census period containing that year, and $n(i)$ the census tract containing the property. The time-varying specification is

$$\ln(p_{it}) \;=\; \delta_t \;+\; \boldsymbol{\theta}_b \, \mathbf{X}_i \, \mathbf{1}[b(t) = b] \;+\; \lambda_{n(i),\, b(t)} \;+\; \varepsilon_{it}, \tag{C.1}$$

where $\ln(p_{it})$ is the log price of property i sold in year t, deflated to 2021 dollars. The term δ_t is a full set of sale-year fixed effects, and $\mathbf{X}_i$ contains the property characteristics, whose coefficients $\boldsymbol{\theta}_b$ vary by census period. We include log floor area, the number of bedrooms, the number of bathrooms, a construction-era dummy, and a strata dummy. The construction-era dummy is a thirteen-category vintage variable built

[6] These criteria include qualified sales, positive areas and prices, and winsorized sales prices at the 1st and 99th percentiles.

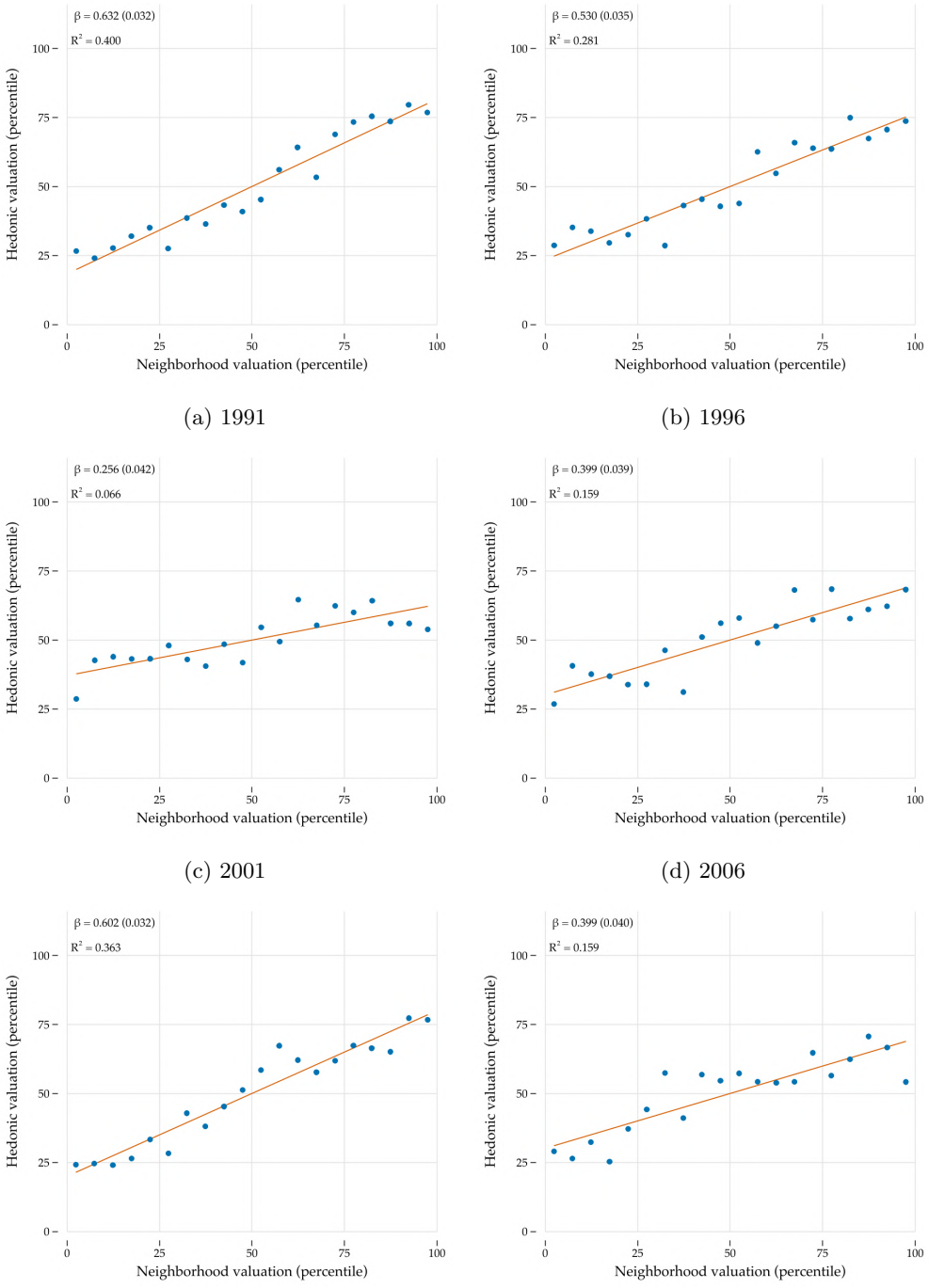
from the recorded year built, with decade bins from the 1920s to the 2020s, a pre-1920 category, and an "Unknown" category for the two sales with missing data. We also include log lot size, stories, a multi-car garage count, a single garage count, a carport count, and a pool dummy, which apply to ground-oriented houses and are undefined or meaningless for strata units.

The term $\lambda_{n,b}$ is the estimated hedonic neighborhood valuation, our parameter of interest. Since both neighborhood valuations are in different units, we compute the percentile distribution for each year for both measures. Thus, we compare how valued each neighborhood is within its corresponding valuation distribution, rather than the levels of each measure.

C.3 Results

Figure C.1 shows the relationship between the two measures of the neighborhood valuations, which we discuss in Section 5.2. The slopes range from 0.26 in 2001 to 0.63 in 1991. We drop the hedonic valuation for 2021 as the data do not cover the full time window.

Figure C.1: Correlation between neighborhood valuations



(a) 1991 (b) 1996

(c) 2001 (d) 2006

(e) 2011 (f) 2016

Notes: The figure shows the relationship between the hedonic neighborhood valuation estimated from property sales data and the neighborhood valuation estimated using migration flows, following the procedure described in Section 5.2, for Vancouver. We transform every valuation to its within-year percentile, so that both measures are in the same units. We aggregate the migration valuations to be independent of immigrant groups, as we do not observe buyers' characteristics in the purchase data. We drop 2021 because the sales data have limited coverage.

D Amenities

In this section, we compare our neighborhood valuation estimated using flows between neighborhoods with neighborhood amenities estimated by Statistics Canada for 2021.

D.1 Construction of the Amenity Measures

Each amenity is built for 2021 census tracts in the Vancouver metropolitan area.

The amenity measures come from the 2022 release of the Spatial Access Measures (Statistics Canada, 2023b). The spatial unit is a dissemination block, and Statistics Canada calculates each measure following Newstead et al. (2024). The amenity values used are: healthcare access, employment access, K12 education access, post-secondary education access, travel time to the closest grocery store, access to cultural venues, and access to sports facilities. All amenities are calculated based on two methods of transportation: by transit at peak hours and by walking, effectively doubling the measures. Newstead et al. (2024) provides more details on how these measures are constructed.

Given that census blocks are more granular than census tracts, we assign each block to a 2021 census tract based on the location of the block's centroid. We then aggregate the measures to the census tract level by taking a population-weighted average of each block's measures. To make measures comparable, we use the percentile of the within-Vancouver distribution of census tracts. Except for travel time to the closest grocery store, higher percentiles mean the tract has a better value of the amenity under study. We also calculate the distance from the tract's centroid to the city center.

D.2 Additional Results

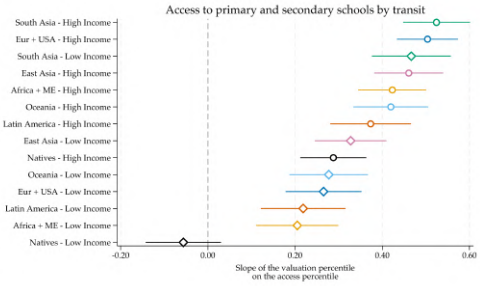
Figures D.1 and D.2 are discussed in Section 5.2. Figure D.3 plots the pooled valuations against the distance to the city center. Figures D.4 and D.5 show the binned relationships between the pooled valuations and access to each amenity, by transit and by walking. Figures D.6 and D.7 compare the 2021 valuations of rich and poor natives with their valuations in earlier census years. Figures D.8 to D.10 compare the valuations of rich and poor households from the same group in each census year. Finally, Figures **??** and **??** map the valuation percentiles in 2016.

Figure D.1: Correlation between amenities and pooled valuations

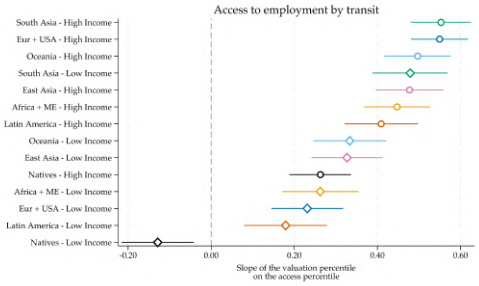


Notes: The figure shows the relationship between the neighborhood amenity and the neighborhood valuation estimated using migration flows, following the procedure described in Section 5.2. The sample only includes neighborhoods in Vancouver in 2021. We transform every valuation to its within-year percentile, so that all measures are in the same units. Each row shows the coefficient of a regression of the pooled neighborhood valuation percentile on the corresponding amenity percentile.

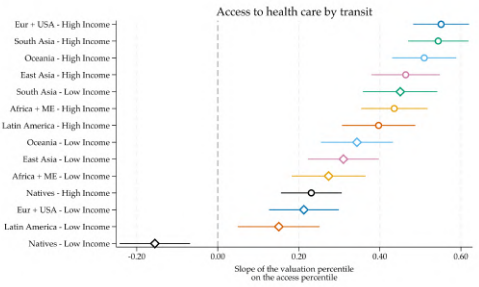
Figure D.2: Correlation between amenities and valuations for different groups



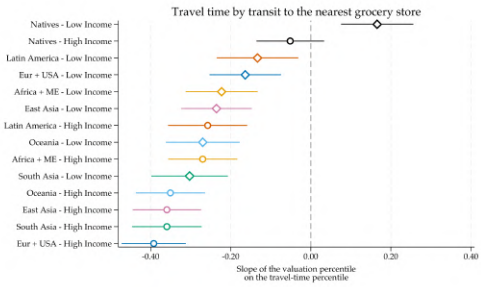
(a) Access to primary and secondary schools



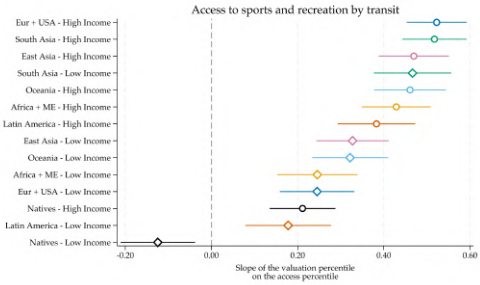
(b) Access to employment



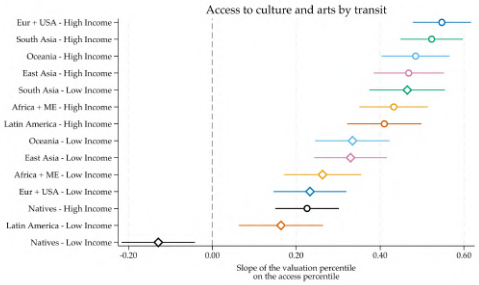
(c) Access to health



(d) Travel time to the closest grocery



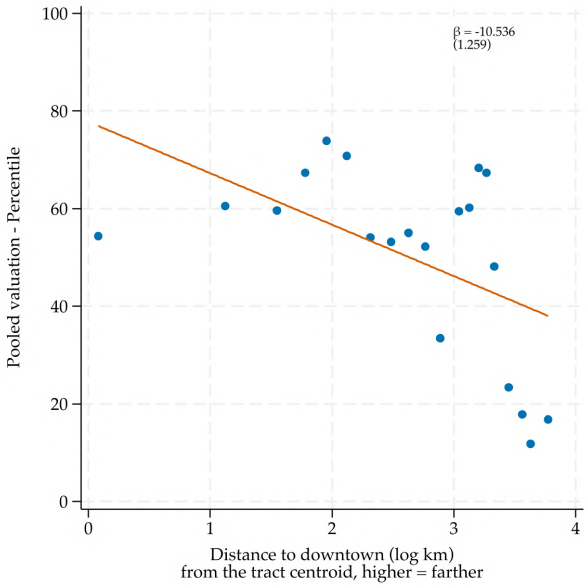
(e) Access to sports and recreational facilities



(f) Access to cultural venues.

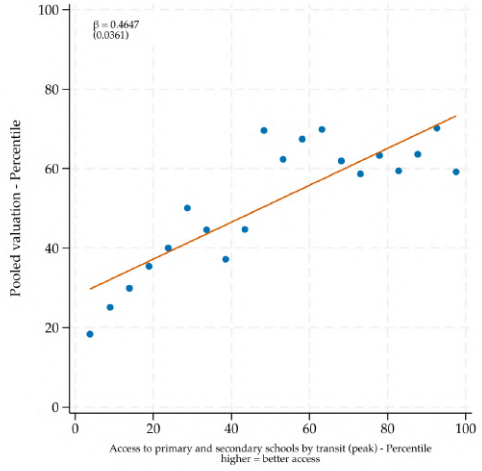
Notes: The figure shows the relationship between the neighborhood amenity and the neighborhood valuation estimated using migration flows, following the procedure described in Section 5.2. The sample only includes neighborhoods in Vancouver in 2021. We transform every valuation to its within-year percentile, so that all measures are in the same units. We measure access by transit in peak hours. Each panel shows the estimated coefficients of regressions of the neighborhood valuation percentile on the amenity percentile. Within each panel, each row corresponds to a different group. Each panel uses a different amenity as the explanatory variable.

Figure D.3: Correlation between pooled valuations and distance to the city center

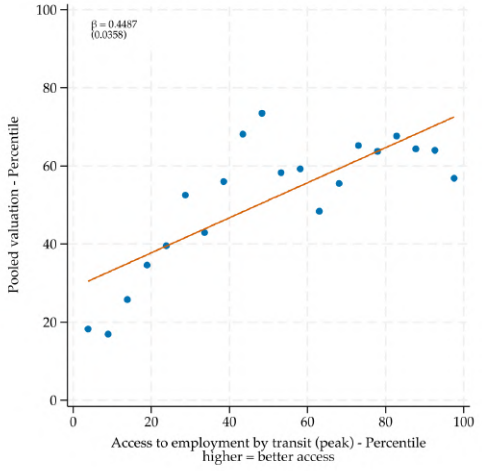


Notes: The figure shows the relationship between the distance from the census tract to the city center and the neighborhood valuation estimated using migration flows, following the procedure described in Section 5.2. The sample only includes neighborhoods in Vancouver in 2021. We transform every valuation to its within-year percentile, so that all measures are in the same units.

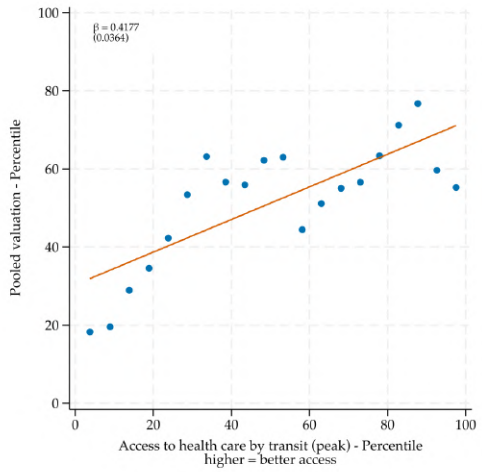
Figure D.4: Correlation between transit access and valuations for the pooled sample



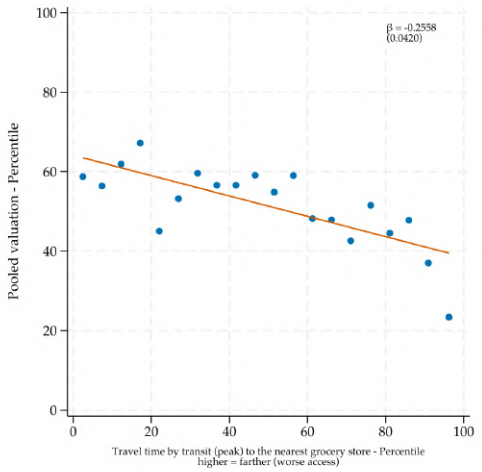
(a) Access to primary and secondary schools



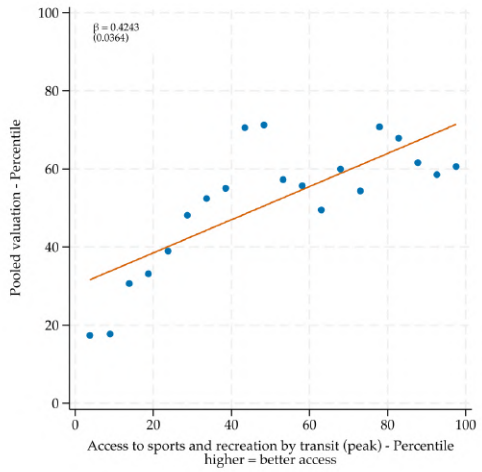
(b) Access to employment



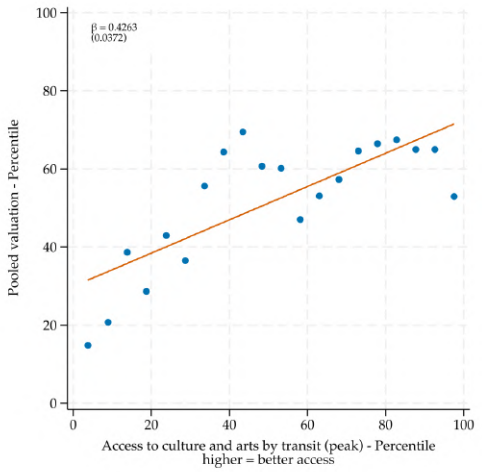
(c) Access to healthcare



(d) Travel time to the closest grocery



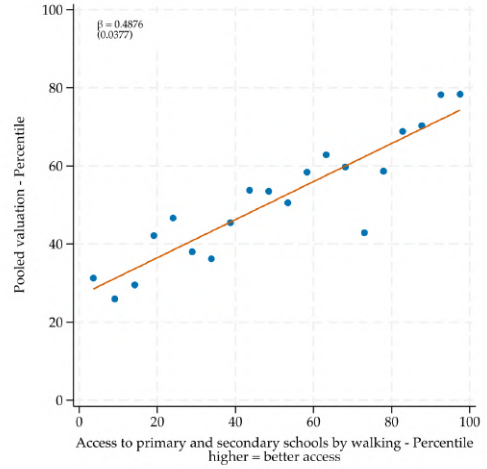
(e) Access to sports and recreational facilities



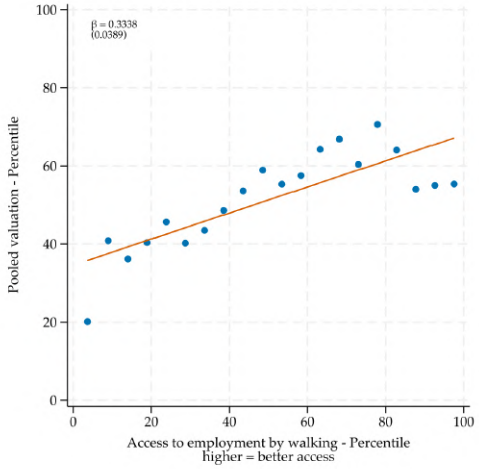
(f) Access to cultural venues.

Notes: The figure shows the relationship between the neighborhood access by transit at peak hours and the neighborhood valuation estimated using migration flows, following the procedure described in Section 5.2. The sample only includes neighborhoods in Vancouver in 2021. We transform every valuation to its within-year percentile, so that all measures are in the same units. We measure access by transit in peak hours. Each panel shows the binscatter of the pooled neighborhood valuation percentile against the amenity percentile. Each panel uses a different amenity.

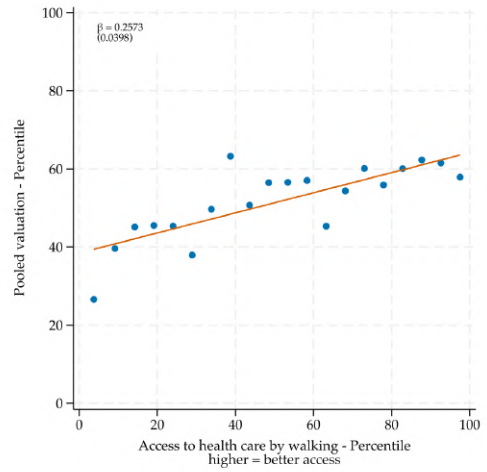
Figure D.5: Correlation between walking access and valuations for the pooled sample



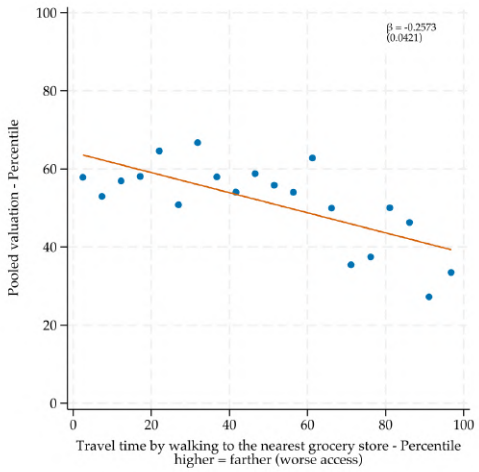
(a) Access to primary and secondary schools



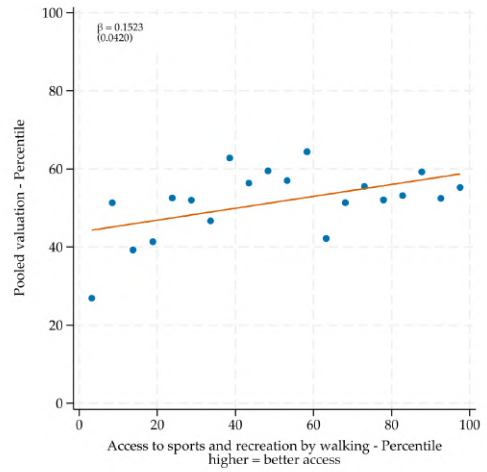
(b) Access to employment



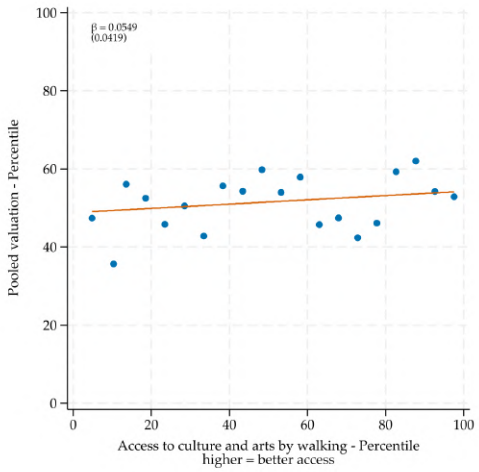
(c) Access to healthcare



(d) Travel time to the closest grocery



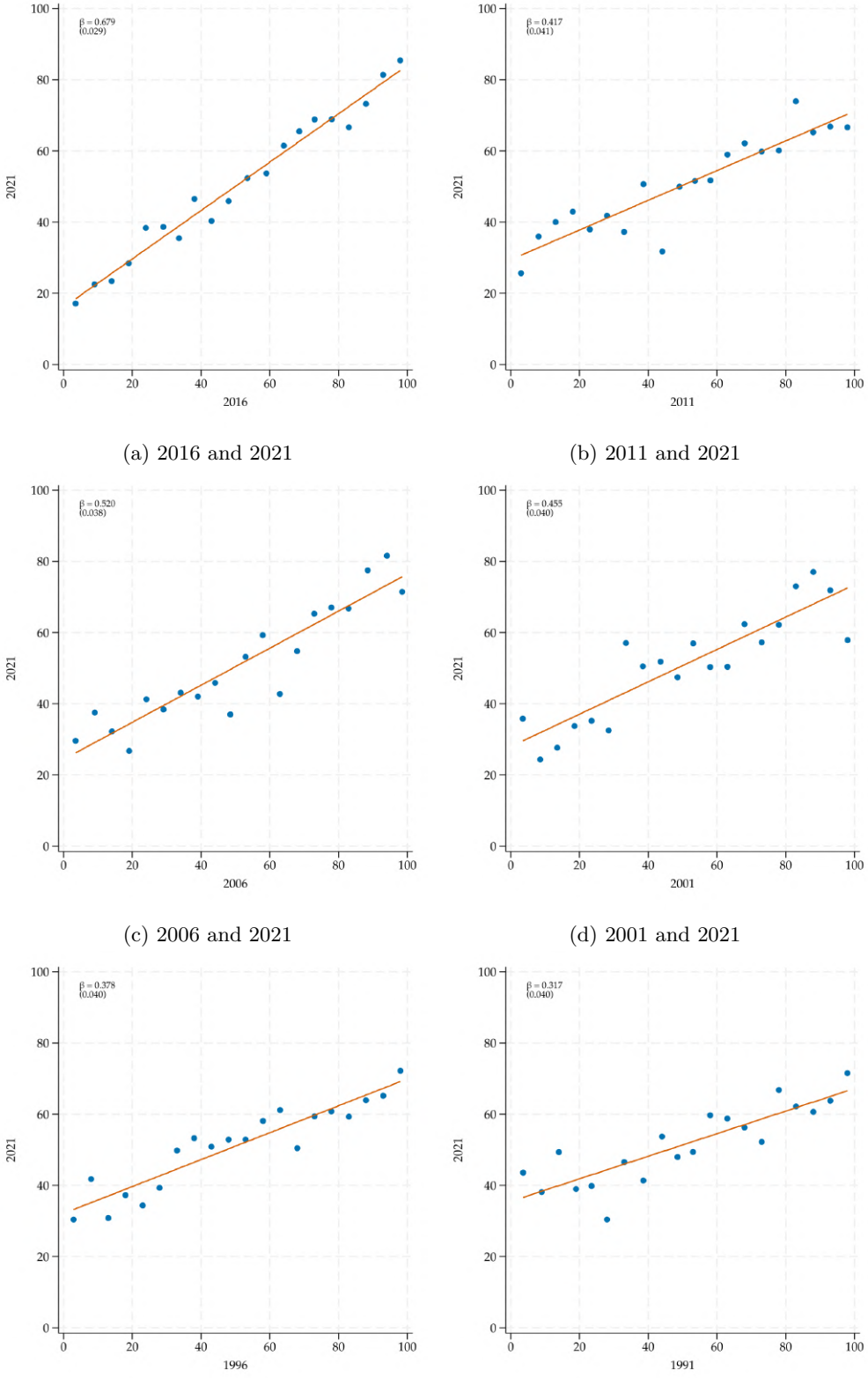
(e) Access to sports and recreational facilities



(f) Access to cultural venues.

Notes: The figure shows the relationship between neighborhood walking access and the neighborhood valuation estimated using migration flows, following the procedure described in Section 5.2. The sample only includes neighborhoods in Vancouver in 2021. We transform every valuation to its within-year percentile, so that all measures are in the same units. We measure access by walking. Each panel shows the binscatter of the pooled neighborhood valuation percentile against the amenity percentile. Each panel uses a different amenity.

Figure D.6: Correlation between neighborhood valuation for Rich Natives across time



(a) 2016 and 2021

(b) 2011 and 2021

(c) 2006 and 2021

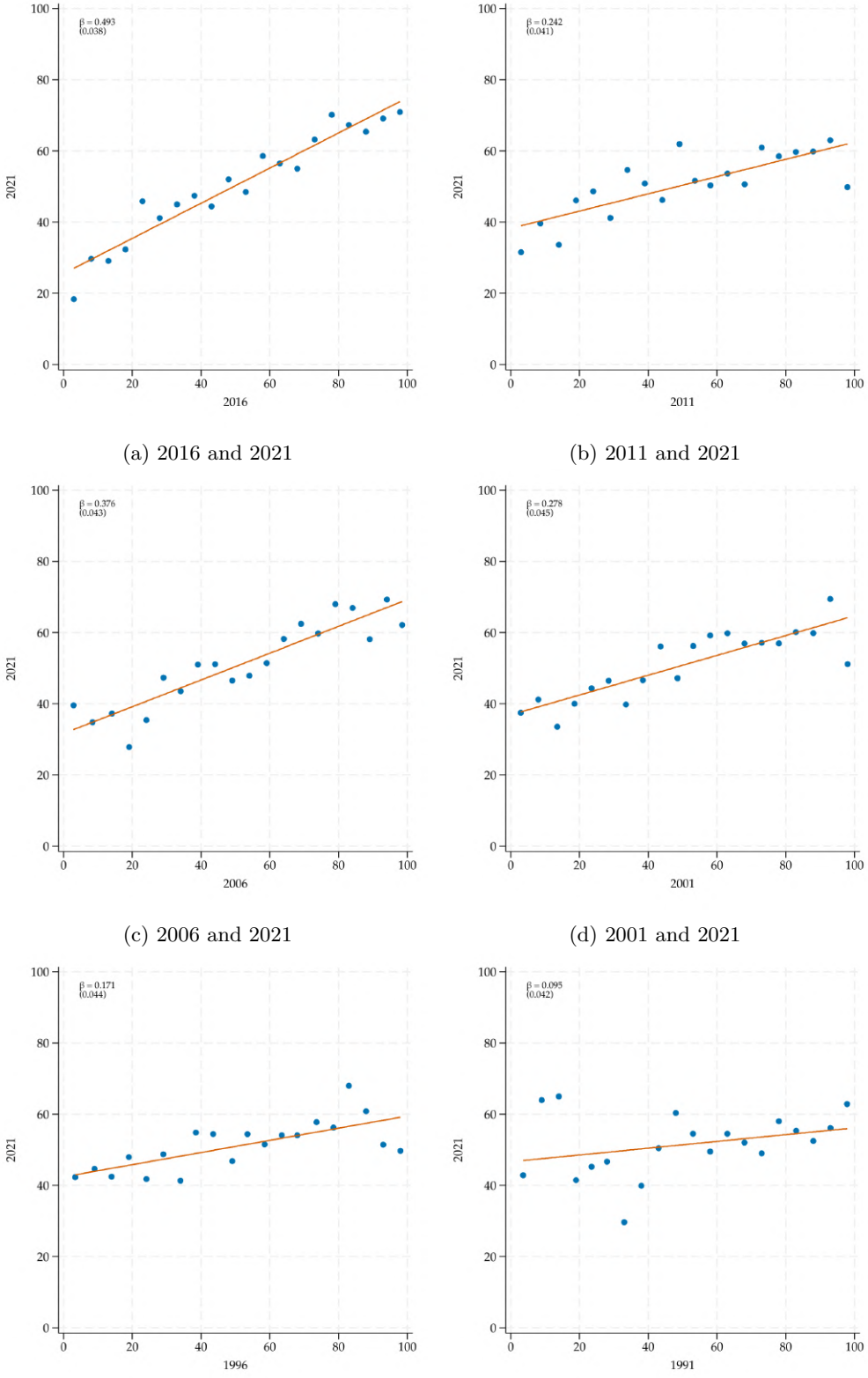
(d) 2001 and 2021

(e) 1996 and 2021

(f) 1991 and 2021

Notes: The figure shows the relationship between the 2021 neighborhood valuation and the valuation in each earlier census year, estimated using migration flows following the procedure described in Section 5.2. The sample only includes valuations for Rich Natives in neighborhoods in Vancouver. We transform every valuation to its within-year percentile, so that all measures are in the same units.

Figure D.7: Correlation between neighborhood valuation for Poor Natives across time



(a) 2016 and 2021

(b) 2011 and 2021

(c) 2006 and 2021

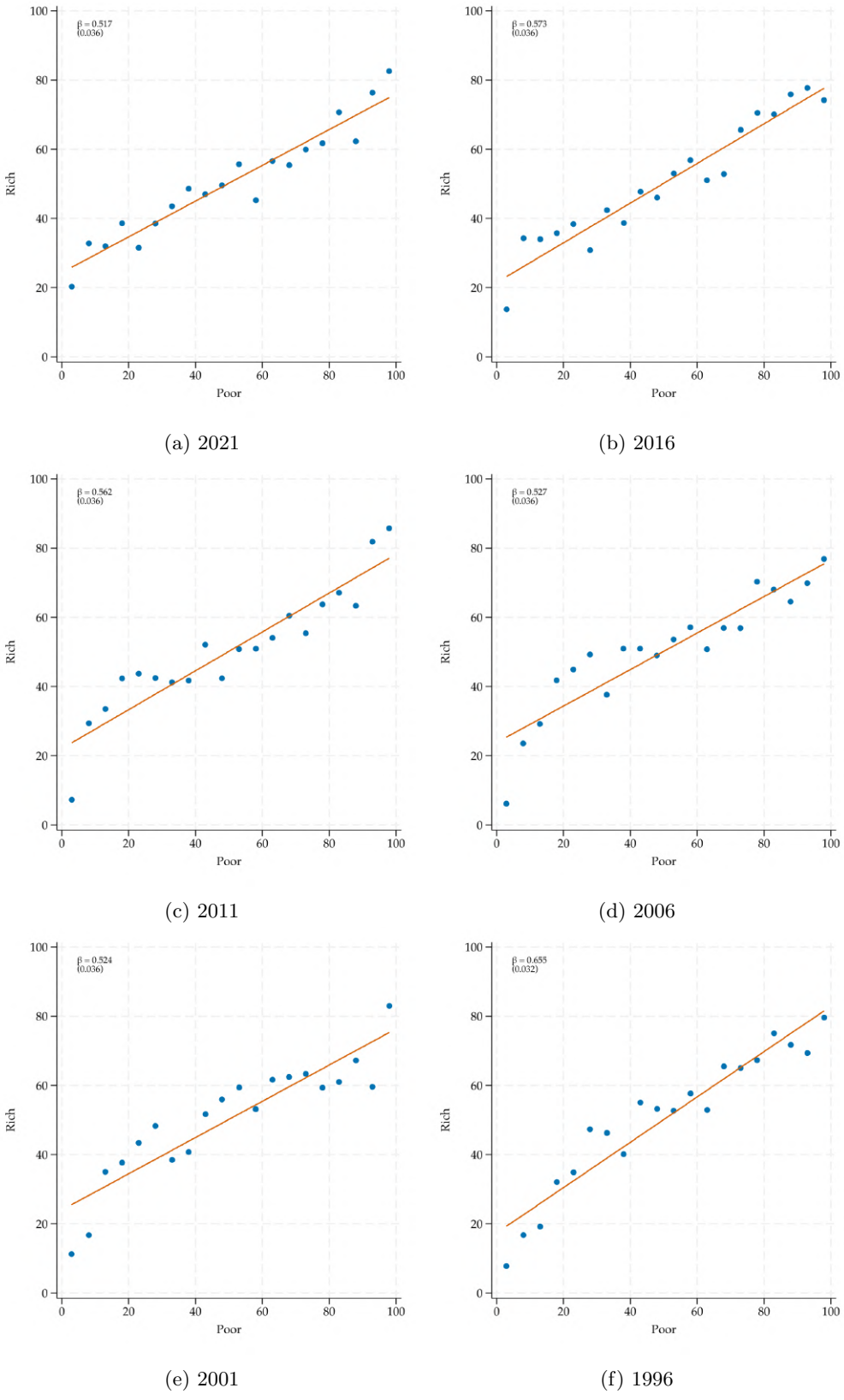
(d) 2001 and 2021

(e) 1996 and 2021

(f) 1991 and 2021

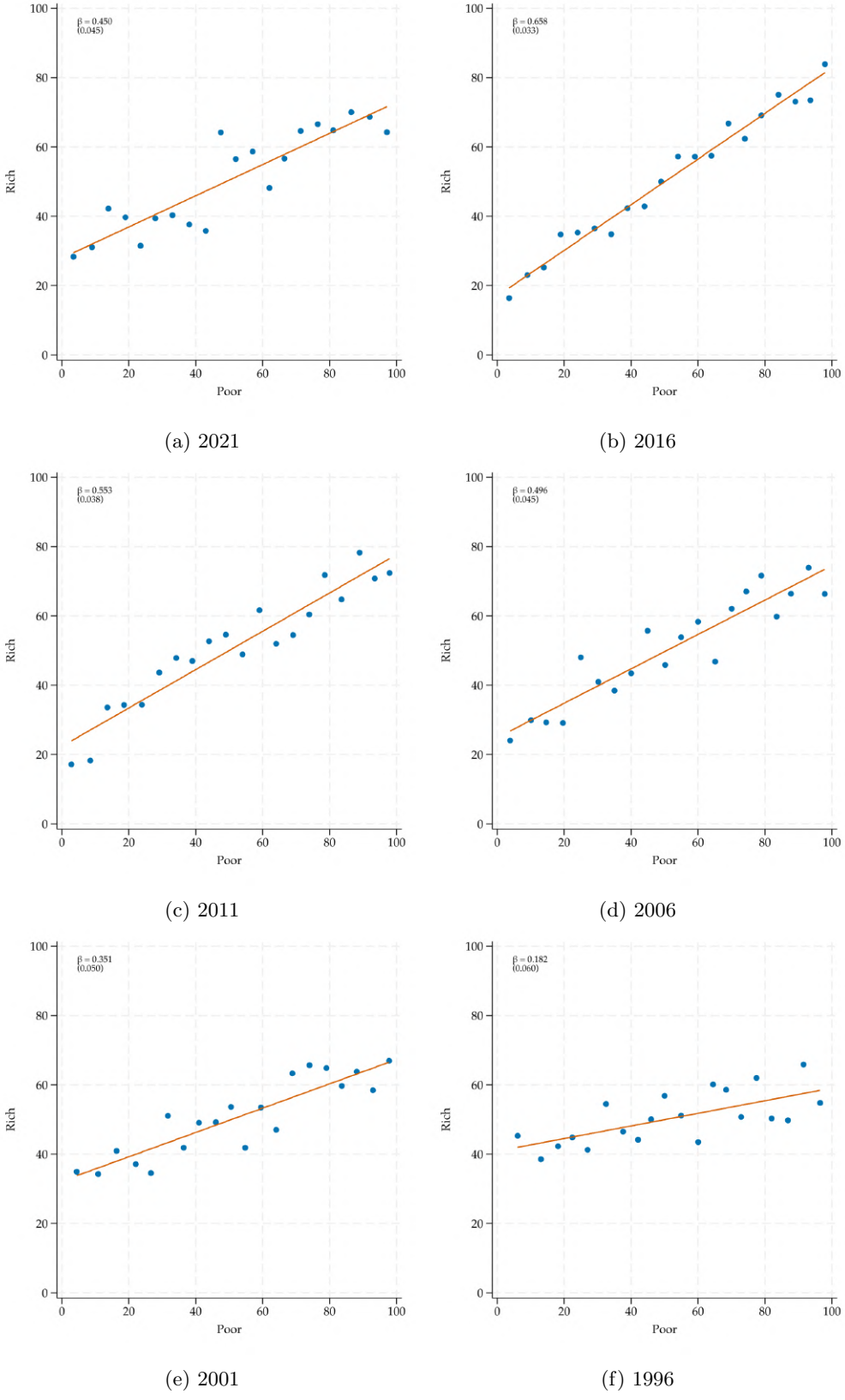
Notes: The figure shows the relationship between the 2021 neighborhood valuation and the valuation in each earlier census year, estimated using migration flows following the procedure described in Section 5.2. The sample only includes valuations for Poor Natives in neighborhoods in Vancouver. We transform every valuation to its within-year percentile, so that all measures are in the same units.

Figure D.8: Correlation between Rich and Poor Natives over time



Notes: The figure shows the relationship between the neighborhood valuations of Rich and Poor natives in the same census year, for each census year from 1996 to 2021, estimated using migration flows following the procedure described in Section 5.2. The sample includes only valuations in Vancouver neighborhoods. We transform every valuation to its within-year percentile, so that all measures are in the same units.

Figure D.9: Correlation between Rich and Poor East Asians over time



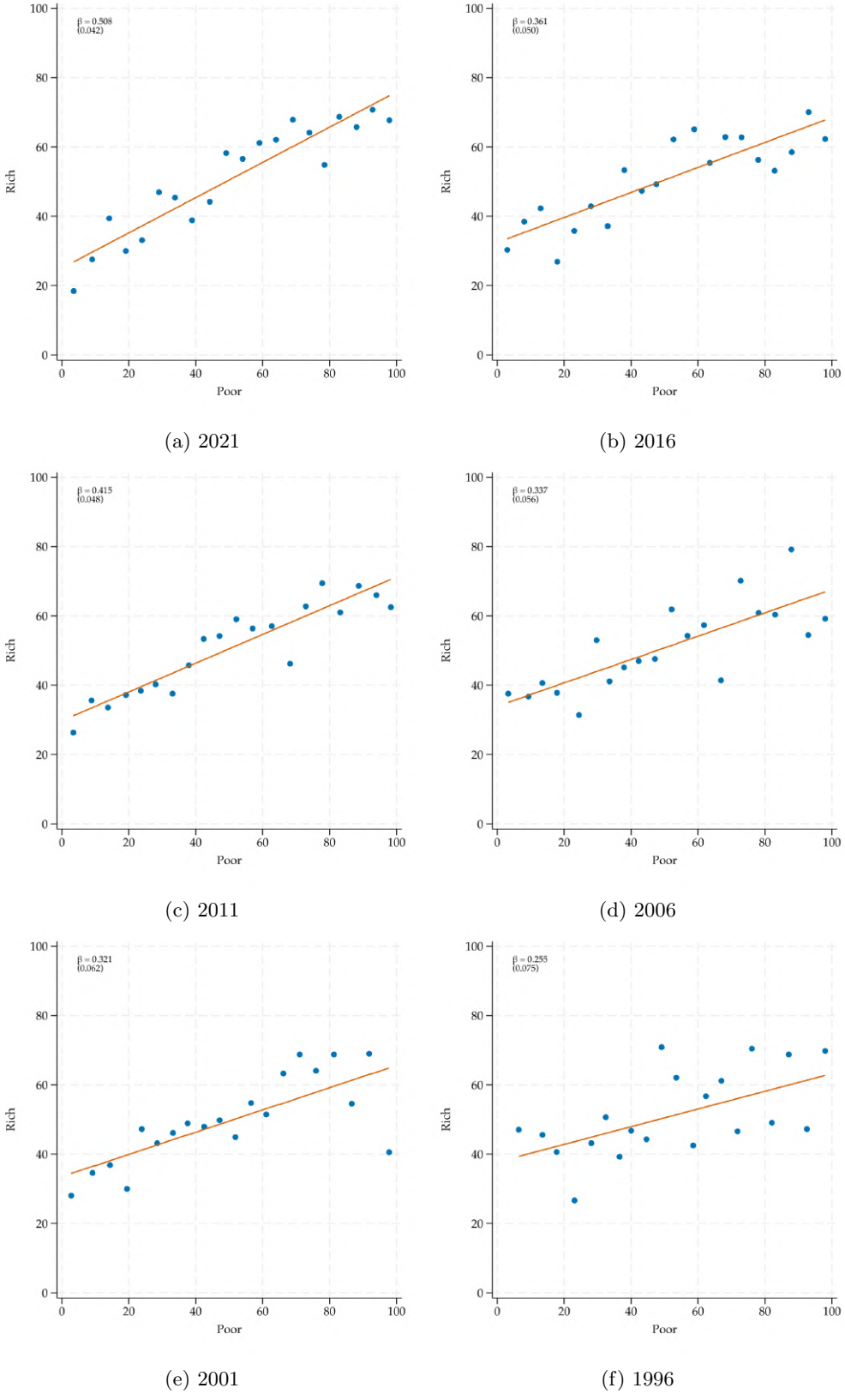
(a) 2021 (b) 2016

(c) 2011 (d) 2006

(e) 2001 (f) 1996

Notes: The figure shows the relationship between the neighborhood valuations of Rich and Poor East Asian residents in the same census year, for each census year from 1996 to 2021, estimated using migration flows following the procedure described in Section 5.2. The sample includes only valuations in Vancouver neighborhoods. We transform every valuation to its within-year percentile, so that all measures are in the same units.

Figure D.10: Correlation between Rich and Poor South Asians over time



(a) 2021 (b) 2016

(c) 2011 (d) 2006

(e) 2001 (f) 1996

Notes: The figure shows the relationship between the neighborhood valuations of Rich and Poor South Asian residents in the same census year, for each census year from 1996 to 2021, estimated using migration flows following the procedure described in Section 5.2. The sample includes only valuations in Vancouver neighborhoods. We transform every valuation to its within-year percentile, so that all measures are in the same units.