The Spatial Distribution of Income in Cities: New Global Evidence and Theory

Peter Deffebach, David Lagakos, Yuhei Miyauchi, Eiji Yamada

October 14, 2025

Abstract

We study how the spatial distribution of income and commuting patterns within cities varies across the development spectrum, drawing on new granular data from 50,000 neighborhoods in 122 cities across 27 developed and less-developed countries. In less-developed cities, poorer residents are more likely to live farther from city centers and in hills and rivers – those with natural amenities but are nonetheless distant from jobs. In developed cities, these patterns are absent or even reversed. Commuting shares also fall more steeply with distance in less-developed countries, reflecting greater commuting frictions that exacerbate inequality in job access. To interpret these findings, we develop a quantitative urban model that incorporates residential and commuting choices with non-homothetic preferences for housing and amenities. We find that the differences in urban spatial income distribution are largely driven by non-homothetic preferences over amenities, higher commuting costs, and more spatially concentrated jobs in less-developed cities. Our model also predicts that as cities grow richer, welfare gains are distributed unequally across households.

Email: peterwd@bu.edu, lagakos@bu.edu, miyauchi@bu.edu, yamada.eiji@jica.go.jp. For helpful comments, we thank David Atkin, Nate Baum-Snow, Olivia Bordeu, Mattias Busso, Juan Pablo Chauvin, Victor Couture, Dave Donaldson, Gilles Duranton, Cecile Gaubert, Ed Glaeser, Nina Harari, Stephen Heblich, Shao-Yu Jheng, Jeff Lin, Ben Marx, Ben Olken, Giacomo Ponzetto, Ewan Rawcliffe, Richard Rogerson, Cezar Santos, Aysegul Sahin, Michael Song, Nick Tsivanidis, Bill Wheaton, and seminar participants at Berkeley, Brown, CUHK, Harvard/MIT, the Inter American Development Bank, Kyoto, Princeton, Toronto, Washington, the World Bank, and various conferences. We thank Japan International Cooperation Agency (JICA) for sharing the data of travel surveys, and Yanbo Pang and Yoshihide Sekimoto for access to the Tokyo travel surveys. For outstanding research assistance, we thank Samantha Agnelli, Daniera Ariefti, Sanghwa Ahn, Lydia Barski, Brayden Binder, Younghoon Chang, Ruifeng Ding, Ryan Duswalt, David Klein, Maria Mittelbach, Isaac Otoo, Shashank Ramachandran, Muhammad Bin Salman, Bhavi Shah, Seungyoo Shin, and Michael Toolan. This project was supported by STEG and the IGC. All potential errors are our own.

1. Introduction

More than half of the world's population now lives in cities. The majority of these urbanites – totaling an estimated 3.7 billion people – reside in less developed nations, in crowded metropolises like Dhaka or Dar es Salaam (United Nations, 2022). For some people, cities serve as the pathway out of poverty, providing reliable access to earning opportunities (Busso, Carrillo, and Chauvin, 2023). Yet many others are stuck in slums, where jobs are few and far away (Marx, Stoker, and Suri, 2013). A growing literature on the urban economics of low- and middle-income countries has sought to better understand how cities can foster job opportunities for wide swaths of the urban population and not just a selected few. However, this research has typically focused on individual cities in the developing world, studying the distributional implications of specific policy changes (see Bryan, Glaeser, and Tsivanidis, 2020; Bryan, Frye, and Morten, 2025, and the references therein). The literature still lacks a comprehensive picture of how cities vary across the development spectrum and how economic development shapes the lives of the urban poor relative to the urban rich.

This paper contributes by building and analyzing a new dataset with granular information on income and commuting covering 50,000 neighborhoods in 122 cities of all levels of income per capita. No existing dataset comprehensively captures neighborhood-level income and commuting patterns in a way that allows cross-city comparisons – particularly in developing countries, where official statistics are often limited. To fill this gap, we construct a harmonized database by combining multiple sources.

Our primary data sources for developing countries are microdata of travel surveys conducted by the Japan International Cooperation Agency (JICA), which they collect as part of urban transportation projects in partner countries. These surveys report residential and workplace locations for an average of 60,000 urban residents per city, along with demographic, employment, and income characteristics. We complement this information with fine-grained administrative census data from selected developing and developed countries. The resulting data cover many of the world's largest Asian and African cities, middle-income megacities such as Lima and São Paulo, and major developed-world cities, including Los Angeles, London, Paris, and Tokyo.

Using this database, we start by documenting several prominent ways in which the residential income distribution within cities varies across the development spectrum. First, we show that in less developed cities, average residential income decreases steadily with distance to city centers. In developed-country cities, in contrast, income-distance gradients are generally flat (in Western Europe) or increasing (in the U.S.). Our focus on distance to the city center is motivated by the long tradition in urban economics positing that more central locations have a higher

concentration of jobs (e.g. Alonso, 1964). In simple terms, our first fact says that in less developed economies, poorer city dwellers are located farther from employment centers on average than richer ones.

Second, we examine how neighborhood incomes within cities vary with the presence of natural amenities. We focus on proximity to hills and waterways, features often associated with higher residential values in developed countries (Lee and Lin, 2018). In developing countries, however, such areas frequently face disadvantages: poor transportation infrastructure that limits job access and inadequate utilities such as sewage systems (McCulloch, Schaelling, Turner, and Kitagawa, 2025). We find that in less developed cities, average incomes are significantly lower in hilly neighborhoods or those near waterways, whereas in developed cities, incomes are flat (in Western Europe) or even increasing (in the U.S.). These sharp cross-country differences likely reflect the interaction between residential sorting by income group and the quality of transportation and residential infrastructure. Later in the paper, we employ a quantitative spatial model to disentangle these mechanisms.

We next examine patterns of commuting and job access, and how they relate to the spatial distribution of residential income documented earlier. We estimate commuting gravity equations for each city, measuring the decay of commuting shares across neighborhoods over bilateral distance, while controlling for residential and workplace fixed effects. We find that less-developed cities exhibit significantly higher commuting frictions. In the U.S., the semi-elasticity of commuting with respect to road distance is approximately 0.1, implying that an additional kilometer reduces commuting by 10 percent, holding workplace attractiveness constant. In Western Europe and Japan, the semi-elasticities are around 0.2, which is somewhat higher than in the U.S., potentially reflecting differences in reliance on private cars versus public transit. In developing countries, the average semi-elasticity rises to 0.35, with a large variation across cities. The estimated semi-elasticities are closely associated with road traffic speeds measured via the Google API (Akbar, Couture, Duranton, and Storeygard, 2023b), while other factors, such as private car penetration, also play a role.

The higher commuting semi-elasticity in less-developed cities suggests that neighborhoods differ markedly in job access within those cities. To explore this point further, we construct a proxy for job access as the distance-discounted sum of workplace fixed effects (Tsivanidis, 2025) and examine its relationship with residential income. In less-developed cities, job access is negatively correlated with neighborhood income. This mirrors our earlier finding that in less-developed cities, neighborhoods located away from the city center or near hills and rivers – often remote from employment hubs – tend to be poorer. In contrast, in developed cities, including the U.S. and Western Europe, the relationship is positive.

These results have important implications for the unequal benefits of urban space. In less-developed countries, poorer residents are more likely to live in neighborhoods with limited access to jobs, while richer residents benefit from central locations with better employment access, potentially reinforcing income inequality. In contrast, in developed countries, wealthier residents tend to live farther away from employment, making room for poorer residents. To interpret these contrasting patterns, uncover their underlying drivers, and assess their welfare consequences, we develop a quantitative spatial model.

In our model, ex-ante heterogeneous households with different earning potentials choose residential locations based on job access, housing costs, and neighborhood amenities. We introduce non-homothetic preferences over housing and amenities, implying that the relative importance of job access, housing, and amenities shifts as household income rises. These preferences generate an equilibrium in which high housing costs in "attractive" neighborhoods crowd out poorer residents (if housing is a "subsistence" good). What makes a neighborhood "attractive" fundamentally depends on the overall income levels of residents. In cities with low average incomes, the primary consideration of households is access to jobs, even for the relatively richer households. As income rises, they start to place a higher value on neighborhood amenities (if amenity is a "luxury good"). As a result, the set of attractive neighborhoods changes with a city's income level. Beyond this mechanism, commuting costs and spatial distribution in productivity, amenities, and housing supply further shape equilibrium residential and commuting patterns.

Our model highlights four potential explanations for the contrasting spatial income distribution between developed and less-developed cities. The first is *non-homothetic preference over amenities*, with richer households placing greater value on features such as open space in suburbs, scenic hills, or waterfront views. Second, *commuting costs tend to be higher in less developed countries*, due in particular to weaker transportation infrastructure and lower rates of private car ownership (Akbar, Couture, Duranton, and Storeygard, 2023a; Akbar et al., 2023b; Tsivanidis, 2025). Third, *jobs may be more centralized in developing cities*, which we model as a steeper decline in workplace productivity with distance from the city center (Baum-Snow, 2020; Davis and Dingel, 2020). Finally, *residential amenity value falls faster away from the city center and in hills and rivers in less developed cities*. This could capture a decline in the supply of residential infrastructure, such as sewage or embankment, or a lack of security and property rights (Harari, 2024; McCulloch et al., 2025; Gertler, Gonzalez-Navarro, Undurraga, and Urrego, 2025).[1]

[1]In the U.S., Glaeser, Kahn, and Rappaport (2008) and Su (2022) argue that higher commuting costs faced by lower-income households lead them to sort into central urban areas. We argue that this mechanism cannot account for the disparity between developed and less-developed cities, as lower-income workers in less-developed cities face even *higher* commuting cost penalties (Section 6.3).

How important is each channel quantitatively? We answer this question in three steps. First, we calibrate the model to match commuting and residential income patterns in U.S. cities. Here, we also estimate preferences for amenities and the degree of non-homotheticity to target observed income sorting across neighborhoods with different housing costs, amenities, and job access. Second, we identify differences in overall productivity, commuting costs, and the spatial distribution of productivity between the U.S. and less-developed cities, with the latter two disciplined by the commuting gravity equations estimated earlier. Third, starting from the U.S.-calibrated model, we counterfactually offset these three differences, with each counterfactual corresponding to one of the first three hypotheses laid out above. Since we lack direct data on neighborhood-level amenities and their valuations, we treat the final hypothesis – differences in amenity gradients across development levels – as a residual explanation.

We find that when lowering the overall city productivity, the residential income premiums in suburban areas and those with natural amenities decline substantially and approach zero. The effect on the suburban premium is intuitive, as poorer households care more about job access, and jobs are less common in suburban areas than central parts of the city. Our data show that areas with natural amenities are also places with relatively poor job access, which means they become relatively less attractive residential locations when overall income levels decline. Increasing commuting costs and changing the concentration further reduce the income premiums and turn them negative, though these effects are smaller in magnitude than those of lowering income alone. Together, these three forces account for 80 percent of the observed income premium gaps in suburban, hilly, and river neighborhoods between U.S. cities and less-developed cities.

While these three forces jointly explain much of the observed gap in spatial income distributions, some residual variation remains. These unexplained patterns likely reflect differences in amenity gradients within cities by development status. Central areas could have greater police protection than neighborhoods on the outskirts of town, or better provision of residential infrastructure like plumbing or electricity. The same could be true of hilly areas, where infrastructure is more expensive to provide, and rivers in less developed cities are almost surely dirtier than their counterparts in richer countries. We leave the task of unpacking these residual variations to future work.

We conclude by examining the distributional welfare implications of economic development for urban residents. As city income rises, high-earning-potential households tend to relocate from central urban areas to suburban neighborhoods with better amenities. This relocation eases housing demand pressures in city centers, moderating rent increases and making these areas more affordable. As a result, residents in urban core areas experience relatively higher welfare gains than those outside. Such heterogeneity does not emerge if we shut down nonhomothetic

preferences, because the uniform increase in labor productivity does not induce any changes in income sorting. Hence, understanding how spatial income distributions vary with income levels and how nonhomothetic preferences shape residential location choices is essential for understanding the welfare implications of urban development across different households.

Our paper contributes to a growing body of work studying how cities differ across the development spectrum. Most cross-country comparisons of urban economic activity either rely on city-level aggregate indicators (Chauvin, Glaeser, Ma, and Tobio, 2017; Jedwab, Loungani, and Yezer, 2021; Lebrand and Kleineberg, 2024), or examine aggregate spatial statistics, such as population density gradients (Henderson and Turner, 2020), building density gradients (Ahlfeldt, Baum-Snow, and Jedwab, 2023; Rosenthal-Kay, 2024), or average road speeds (Akbar et al., 2023a,b). Much has been learned as well from detailed analyses of individual cities in developing countries.[2] Our study is closely related to those focused on cross-city comparisons of internal city structure in developing countries, such as Harari (2024) and Adukia, Asher, Jha, Novosad, and Tan (2022), who study income segregation and public access and public goods provision within Brazil and India, respectively, and Dingel, Miscio, and Davis (2021) who use data from Brazil and China to show that residents living closer to city centers are more skilled on average. We contribute to this literature by building and analyzing granular information on income and commuting across 122 cities in 27 countries, covering a wide range of income levels.

We also contribute to the empirical and theoretical literature on the spatial distribution of income within cities. Traditionally, this literature has predominantly focused on the U.S. cities. This work has sought to explain U.S. patterns using non-homotheticities in demand for housing or land (Alonso, 1964; Becker, 1965; Margo, 1992; Hoelzlein, 2023; Finlay and Williams, 2025; Couture, Gaubert, Handbury, and Hurst, 2024), transportation infrastructure (Glaeser et al., 2008; Su, 2022), and natural amenities (Lee and Lin, 2018). A smaller strand of research has highlighted differences between U.S. cities and specific cities in other developed countries, attributing them to other amenities such as restaurants and bars concentrated in central areas (Brueckner, Thisse, and Zenou, 1999 for Paris; Almagro and Domínguez-Iino, 2025 for Amsterdam; Tabuchi, 2019 for Tokyo). By contrast, little has been documented or explained about the patterns in less-developed cities. Our main contribution is to adopt a cross-country perspective: we document how these patterns manifest in less-developed economies and provide an explanation for why they differ.

[2]See, for example, Tsivanidis (2025), Zárate (2024), Balboni, Bryan, Morten, O'Connor, and Siddiqi (2025), Khanna, Nyshadham, Ramos-Menchelli, Tamayo, and Tiew (2023), and Bordeu (2023) for transportation infrastructure in Bogota, Mexico City, Dar es Salaam, Medellin, and Santiago, respectively; Michaels, Nigmatulina, Rauch, Regan, Baruah, and Dahlstrand (2021) and Franklin, Imbert, Abebe, and Mejia-Mantilla (2024) for residential infrastructure and urban public works program in Dar es Salaam; and Harari and Wong (2024) and Gechter and Tsivanidis (2023) for slum upgrading interventions in Jakarta and Mumbai, respectively.

While our central focus is the comparison between developed and less-developed cities, we also present systematic evidence on developed cities outside the U.S., rather than relying on a few specific megacities, such as Paris or Tokyo. On average, Western European cities exhibit flat, rather than negative, income gradients with distance to the city center, as well as with hills and rivers. We show that the differences between U.S. cities and other developed cities in Western Europe and Japan stem from differences in amenity dimensions beyond the suburban and natural amenities emphasized in the U.S. context, and to a lesser extent, differences in commuting costs and spatial productivity distribution.

Although our analysis is cross-sectional comparisons across cities, our findings also speak to the literature on how the spatial organization of economic activity evolves with development. Empirically, our result that residential income declines monotonically with distance to the city center in less-developed cities parallels the evidence in Lee and Lin (2018), who show that U.S. cities exhibited this pattern in 1880, but that it reversed after 1930. Theoretically, our paper relates to work on the interaction between structural transformation of industries – from agriculture, manufacturing, and services – and spatial population reallocation across cities (Eckert and Peters, 2022; Chatterjee, Giannone, and Kuno, 2021; Bohr, Mestieri, and Robert-Nicoud, 2024) or within them (Coeurdacier, Oswald, and Teignier, 2025). While related in its focus on the non-homothetic preferences, our focus is on housing and amenities, and how they shape spatial income distribution within cities.

2. Data

This section outlines how we integrate diverse data sources to build a comprehensive database on household income and commuting flows from 50,000 neighborhoods in 122 cities across 27 countries.

2.1. Travel Surveys from Developing Countries

Our primary data source for developing countries is microdata of travel surveys from 27 cities in 21 developing countries conducted by the Japan International Cooperation Agency (JICA), the official aid agency of Japan. JICA collects these data as part of urban transportation projects in partner countries. These surveys gather detailed information on residential and workplace locations, demographic characteristics (e.g., age, household size), income, employment status, and daily travel activities, including trip timing, geolocation, purpose, and transportation mode. In developed countries, such surveys are usually carried out at regular intervals in major metropolitan areas. In developing countries, they are often conducted on an ad hoc basis, typically

in preparation for major infrastructure projects or city master plans, and often with support from international aid agencies such as JICA.

Panel (a) of Table 1 provides a list of the cities included in this dataset. The coverage spans multiple continents, including three cities in Latin America (Lima, Managua, and Belém), four cities in South Asia (e.g., Dhaka and Karachi), 11 in East Asia (e.g., Hanoi and Vientiane), one in Eastern Europe (Bucharest), two in the Middle East (Cairo and Damascus), and five in Africa (e.g., Nairobi and Maputo).

The surveys were conducted between 1996 and 2018 and vary in terms of questionnaire design, sampling strategy, and local implementation. Sample sizes range from 5,000 to 300,000 respondents, with an average of 60,000 per city.

Our travel surveys are particularly well-suited to our analysis due to their fine spatial resolution. Each survey divides the city into a large number of neighborhoods, or "survey zones," and records respondents' residential and workplace locations, as well as the origins and destinations of daily trips, at this neighborhood level. In many cases, the survey zones are only available in the form of non-georeferenced maps (i.e., image files). To address this, we manually geo-coded the survey maps, as illustrated in Appendix Figure A.3. On average, there are 193 survey zones per city, with each zone covering approximately 8 square kilometers.

All surveys include household-level income information. In most cities, respondents report their household's total income, either as a continuous value or within finely disaggregated bins. In three cities (Bucharest, Dhaka, and Managua), the surveys do not directly ask household income, but instead collect individual income for each household member, which we aggregate to construct household-level income.

Spatially disaggregated income data are rarely available for cities in developing countries. As such, our dataset represents the first comprehensive effort to measure neighborhood-level income across a broad set of cities in developing countries. Nevertheless, concerns may arise regarding data accuracy, particularly since the travel surveys are based on surveys rather than administrative data based on complete household enumeration. To assess the validity of our income measures, we examine the case of Belém, Brazil, the only city in our sample for which neighborhood-level income from a national census is publicly available (see Section 2.2). In Appendix A.2, we show that the travel survey data closely align with the census data in both the overall spatial income distribution and the gradient of income with respect to distance from the city center.

We also extract households' commuting information from these travel surveys. For each respondent, the survey typically records whether the individual is employed (either as a wage

Table 1: List of Cities Covered in Our Analysis

	L. America	Asia, E. Europe	Africa, M. East
Number of Cities	3	16	7
Number of Countries	3	12	6
Avg Number of Respondents	70469	70965	46113
Avg Number of Neighborhoods	190	220	178
List of Cities	Belem (00), Lima (03), Managua (98)	Bucharest (98), Cebu (14), Chengdu (00), Colombo (13), Da Nang (08), Dhaka (14), Hanoi (05), Ho Chi Minh (14), Jakarta (18), Karachi (11), Kuala Lumpur (99), Lahore (10), Manila (96), Phnom Penh (12), Viang Chan (07), Yangon (13)	Abidjan (13), Cairo (01), Damascus (98), Dar es Salaam (07), Kinshasa (18), Mombasa (15), Nairobi (13)

(a) Less developed cities surveyed by JICA, by region

	USA	W. Europe, Japan	L. America	Asia, E. Europe	Africa, M. East
Number of Cities	48	24	27	16	7
... with Commuting Flows	48	17	3	16	7
Number of Countries	1	4	3	12	6
... with Commuting Flows	1	3	3	12	6
Total Number of Neighborhoods	27579	13170	4215	3527	1246
List of Countries	United States	France, Japan, Spain, United Kingdom	Brazil, Nicaragua, Peru	Bangladesh, Cambodia, China, Indonesia, Lao People's DR, Malaysia, Myanmar, Pakistan, Philippines, Romania, Sri Lanka, Viet Nam	Côte d'Ivoire, D.R. of the Congo, Egypt, Kenya, Syrian Arab Republic, U.R. of Tanzania: Mainland

(b) All cities in neighborhood-level income dataset, by region

Note: List of cities included in our data set. See Appendix Table A.1, A.2, A.3 for the characteristics of each city in our JICA survey data; Table A.4 for the list of cities other than JICA surveys; and Figure A.1 for a map.

worker or self-employed) and, if so, their workplace location, coded at the survey zone level. In seven cities, the surveys do not directly ask about work locations. In those cases, we rely on the travel activity module, which documents the time, location, and purpose of each trip. We infer workplace locations by identifying trips made for the purpose of going to work. Using these data,

we construct origin-destination commuting flows between survey zones within each city.

2.2. Additional Data on Income, Commuting Flows, and Rents

Residential Neighborhood Income from U.S., U.K., France, Spain, and Brazil We collect average residential neighborhood income from census and tax data for four developed countries (U.S., Spain, France, U.K.) and a less-developed country (Brazil). Information from the U.S. stems from the 2015-2019 aggregated American Community Survey at the level of census tract. For the U.K., we obtain average income at the *small area* level from the Office of Tax Statistics in 2018. For France, income is derived from tax returns in 2016 at the level of *communes*, with further disaggregation into *arrondissements* within some metropolitan areas (e.g., Paris).[3] In Spain, average neighborhood income derived from tax information is available at the *sección*-level in 2019. In Brazil, average neighborhood income from the 2010 census is available at the *bairros*-level.[4]

Travel Survey from Tokyo, Japan For Tokyo, Japan, we have access to the microdata of the 2018 Tokyo Person Trip Survey (Tokyo Metropolitan Area Transportation Planning Council, 2018). The data shares a similar structure to the JICA surveys discussed above – an individual-level survey reporting household income, demographic information, discrete neighborhoods, home and work location, and trips throughout the day.

Bilateral Commuting Flows in U.S., U.K., France We obtain commuting flow information in the U.S. from the Census Transportation Planning Packages (CTPP) for the years 2012-2016, which are constructed using the American Community Survey (ACS) data on usual residence and workplaces. The CTPP data reports the aggregate number of workers living and working in any given pair of census tracks.[5] We use bilateral commuting flows from the U.K. derived from censuses at the *small area* level, and from France at the *communes/arrondissements* level, with the same level of disaggregation as the income data. There are no publicly available commuting flow data from Spain and Brazil (except for Belem, where the JICA travel survey is available).

[3] We use median neighborhood income, instead of mean income, for France, as the latter is not publicly available. While income information is available at a finer disaggregation at *IRIS* level, we use *communes/arrondissements* because of the availability of commuting flow data.

[4] In some parts of the country where *bairros* are not defined, we use *subdistritos*, a coarser disaggregation. While income information is available at a finer disaggregation at *setores* level (about one-tenth of the area size of U.S. census tract), we use *bairros/subdistritos* because of the comparability of neighborhood sizes with other countries.

[5] An alternative commuting dataset in the U.S. is the Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics (LODES). This dataset is constructed from the Quarterly Census of Employment and Wages, which covers all firms with paid employees and are subject to unemployment insurance laws. While LODES provides a near-census coverage of paid workers with their precise residential locations, it excludes the self-employed and imputes worksite locations for multi-establishment firms. Due to these differences, LODES yields commuting semi-elasticities that are roughly two-thirds the magnitude of those estimated from the CTPP and from travel surveys in Chicago and San Francisco (Appendix A.3; see also Spear, 2011).

Housing Rents in U.S. We obtain median housing rents for U.S. cities from the 2015-2019 aggregated American Community Survey (ACS) at the level of census tract. We use this information for calibrating our quantitative model in Section 6.

2.3. Additional Geographic Data and Definitions

City Boundaries To create a consistent definition of city boundaries for all cities and exclude neighborhoods that lie in rural areas, we use the World Settlement Footprint's "Built Up Areas" data set (Florczyk et al., 2019). This data set uses satellite images to categorize land use at a fine-grained spatial level and defines cities as large contiguous swaths of land which are "built-up," i.e. full of buildings, roads, and pavement, in contrast to agriculture or forest cover. Each city is defined by a geographically contiguous built-up area. This definition of city puts together different municipalities within a broad metropolitan area; for example, New York City extends New Brunswick to the south and White Plains to the north, and London extends out to Heathrow Airport to the west. In our analysis, we include all cities with populations above 400,000 people.

City Centers and Suburban Areas We define city centers using coordinates from OpenStreetMap (OSM), an open-source collaborative mapping platform of the world. Contributors typically assign city center locations based on prominent landmarks such as city halls or central plazas. Although this approach is heuristic, it aligns well with intuitive notions of a city center.[6] As shown in Figure 1, these locations coincide with the highest net commuter densities (in-commuters minus out-commuters per unit area) in cities like Los Angeles and Lima. To address potential measurement error of the exact city center locations or the presence of polycentric structures, we also analyze broader patterns between suburban areas and others, defining *suburban areas* as the neighborhoods comprising 50 percent of the population living farthest from the city center.

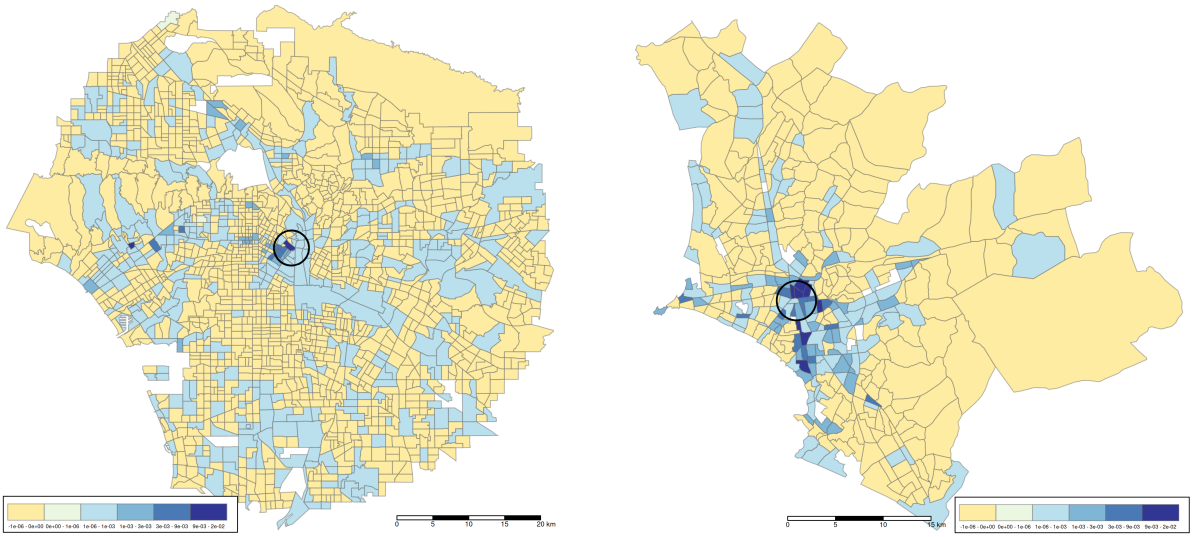
Bilateral Travel Distance To calculate the distance between pairs of neighborhoods in our pairwise commuting flow data set, we use the Open Source Routing Machine (OSRM), an open-source algorithm for finding the shortest path between two locations along OSM's road network (Luxen and Vetter, 2011).

Hills We classify a neighborhood as *hilly* if its average slope (average change in elevation across four adjacent grid cells) exceeds 5 degrees, based on 30m×30m elevation data from Amazon Web Services Terrain Tiles (Larrick, Tian, Rogers, Acosta, and Shen, 2020).[7]

[6] For example, Cebu City's center is defined at the city hall, while London's is at Trafalgar Square. As an additional validation, for the 54 cities with data on "prime locations" (Ahlfeldt, Albers, and Behrens, 2020), we find that in 75 percent of cases the prime locations are within 2 kilometers of our city center definitions.

[7] While Lee and Lin (2018) uses a 15-degree threshold, we adopt a lower cutoff to capture a broader set of moderately sloped areas, reflecting the generally less steep terrain in many cities worldwide.

Figure 1: Net Commuting and City Centers in Los Angeles and Lima



(a) Net commuter density (Los Angeles) (b) Net commuter density (Lima)

Note: The figures show the density of net commuters (total in-commutes minus out-commutes) for each neighborhood in Los Angeles and Lima. Darker blue color indicates a higher net-commute density. Yellow color (the bottom category) indicates negative values for net-commute. The circle indicates 2 kilometers from the city center.

Rivers We identify river proximity using the HydroSHEDS dataset (Lehner and Grill, 2013), which maps global water flows based on topography and rainfall. Neighborhoods are classified as near a river if any part lies within 100 meters of a riverbank, considering only rivers with an average flow exceeding 1.3 cubic meters per second.

Neighborhood Population We construct a standardized measure of population size for each neighborhood using the 2015 LandScan population distributions (Bright, Rose, and Urban, 2016) for all cities that are based on travel surveys. Population estimates from the U.S., U.K., Spain, France, and Brazil are derived from the underlying administrative data source directly.

2.4. Final Data Sets

We define *developed* cities as those in the U.S., Spain, France, U.K., and Japan. We define *less-developed* cities as those surveyed by JICA and Brazil. The poorest country covered in our data is the Democratic Republic of the Congo (for the city of Kinshasa), with a GDP per capita of 1,020 USD. To gauge further heterogeneity within the developed and less-developed cities, we often separately report estimates by *regions*, constituting U.S. ("USA") and Western Europe and Japan ("W. Europe, Japan") for developed cities, and Latin America ("L. America"), Asia and Eastern

Europe ("Asia, E. Europe"), and Africa and Middle East ("Africa, M. East") for less-developed cities.

Our final neighborhood-level income dataset includes around 50,000 neighborhoods across 122 cities in 27 countries. 72 are classified as developed (48 from "USA" and 24 from "W. Europe, Japan") and 55 are classified as less-developed (27 from "L. America," 16 from "Asia, E. Europe," and 7 from "Africa, M. East"). For bilateral commuting flows, we have data at the same level of disaggregation, except for all cities in Spain and all cities in Brazil, other than Belém. Panel (b) of Table 1 lists all the cities in our analysis. Appendix Figure A.1 shows the map of the world with locations of all cities in our data.

3. Spatial Distribution of Income

This section documents several prominent ways in which the spatial distribution of income within cities varies across regions and between developed and less-developed cities.

3.1. A First Look from Examples: Los Angeles and Lima

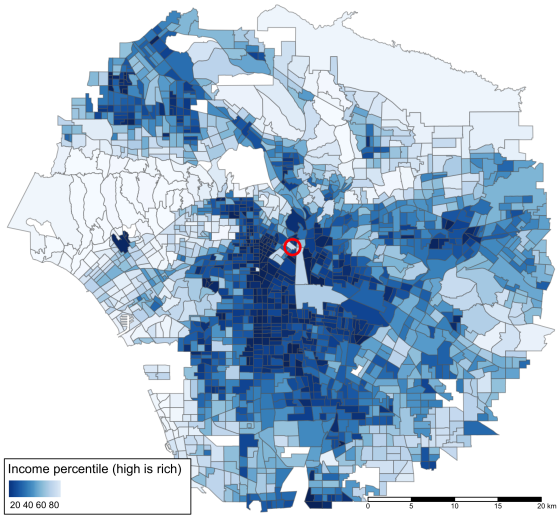
Before proceeding to the full statistical analysis, we begin with an illustrative comparison between two cities: Los Angeles, U.S., and Lima, Peru.

Figure 2 illustrates income distribution and geographic features for Los Angeles and Lima. Panels (a) and (b) show the average residential income by neighborhood, measured as each neighborhood's percentile rank within the city, with lighter colors corresponding to higher income levels. The red circle in each panel marks the city center. Panels (c) and (d) highlight neighborhoods that are hilly or located near major waterways.

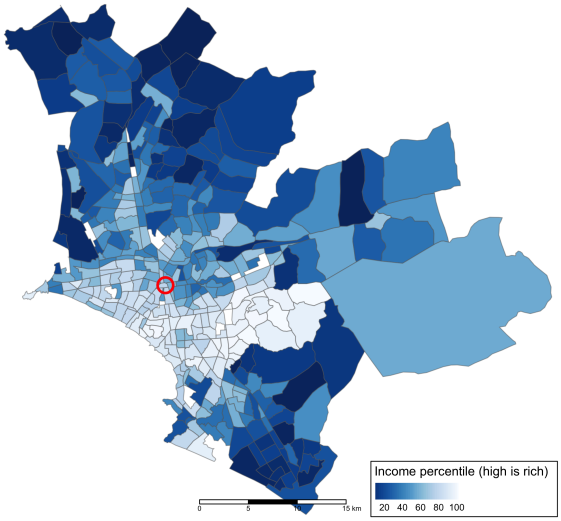
Focusing first on Panels (a) and (b), the two cities display starkly contrasting relationships between average income and distance to the city center. In Los Angeles, lower-income neighborhoods surround the city center, with the exception of a small cluster of higher-income blocks at the core. Moving outward, particularly toward the north (Pasadena) and west (Santa Monica), average income tends to rise. In contrast, Lima exhibits the opposite pattern: neighborhoods near the city center are generally wealthier, and income declines with distance from the center.

Panels (c) and (d) reveal similarly contrasting patterns between income and hilly areas. In Los Angeles, hilly areas such as Laurel Canyon and Beverly Hills – located northwest of the city center – are associated with high incomes. In Lima, by contrast, Los Olivos, a middle-income area nestled in a valley northwest of the center, is surrounded by poorer hillside neighborhoods

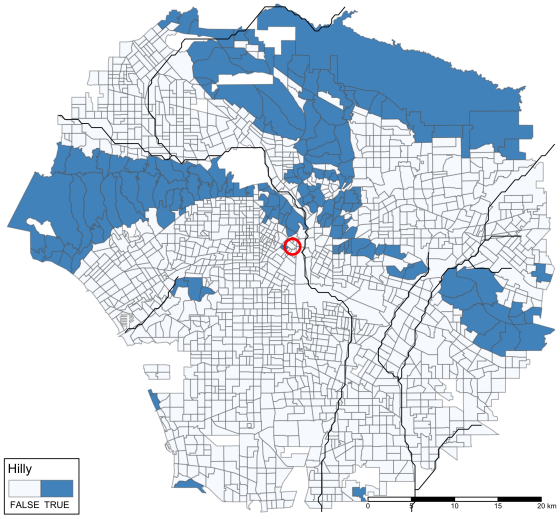
Figure 2: Residential Income and Hilly Areas in Los Angeles and Lima



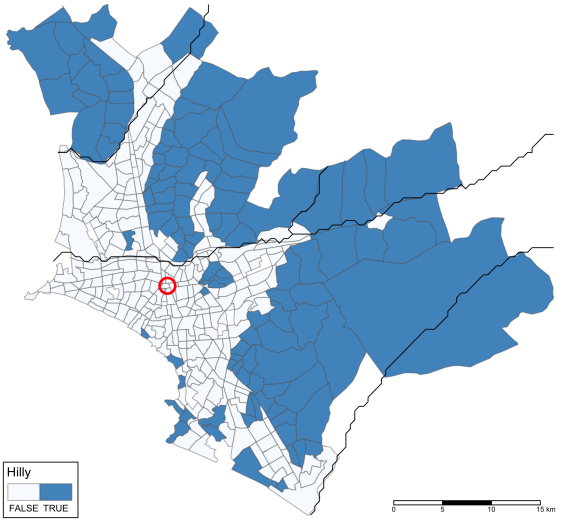
(a) Residential Income (Los Angeles)



(b) Residential Income (Lima)



(c) Hills and Rivers (Los Angeles)



(d) Hills and Rivers (Lima)

Note: Panels (a) and (b) show the average residential income by neighborhood, measured as each neighborhood's percentile rank within the city, with lighter colors corresponding to higher income levels. Figures (c) and (d) show a binary measure for hilliness (blue is hilly) along with the path of waterways in black. The red circle in each panel marks the city center. Neighborhoods further than 30km from the city center are omitted.

on both sides.

In what follows, we show that these patterns reflect broader, systematic differences in spatial

income distribution between developed and less-developed cities, extending beyond the specific cases of Los Angeles and Lima.

3.2. Residential Income and Distance to City Center

We first focus on the residential income and distance to the city center. Figure 3 shows the relationship between distance from the city center and average neighborhood residential income percentiles for each region up to 25 kilometers from the city center. Each light line represents a single city, while averages are highlighted in bold.

Panel (a) shows that in the U.S. cities, income exhibits a modest U-shaped pattern with respect to distance from the city center. On average, neighborhoods exactly at the city center are slightly below the 50th percentile. Income declines up to 4 kilometers from the center, reaching a low point at around the 40th percentile. Beyond this point, income gradually increases toward the urban periphery. Overall, this pattern reflects a positive income gradient from the inner suburbs to the outer edges of the city.

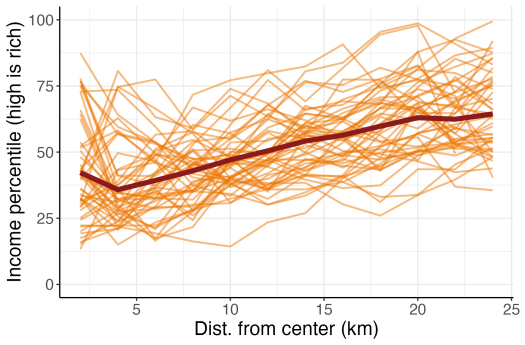
In Panel (b), we show the patterns of other developed cities from Western Europe and Japan. We find that, on average, the income profiles are mostly flat. Some specific megacities, such as Tokyo or Paris, exhibit downward-sloping patterns (see Appendix Figure B.3), but these cases are not representative of the broader set of European cities.

In contrast, in Panels (c)-(e), we find a strong, monotonic decline in income with distance from the city center in less-developed cities. Across all regions, central neighborhoods average around the 60th percentile in income, but this steadily falls away from the center, indicating a pronounced negative gradient.

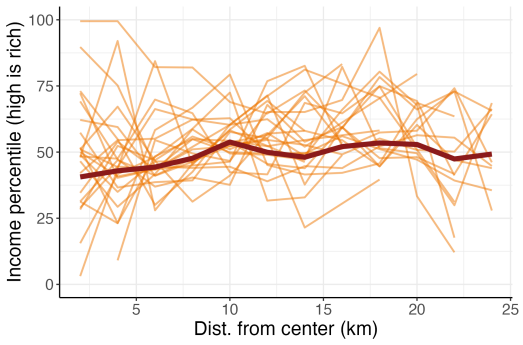
Figure 4 further examines cross-city variation in the relationship between income and distance to the city center. For each city, we compute the "suburban–urban income gap," defined as the difference in average income percentiles between suburban neighborhoods (neighborhoods housing the 50 percent of the population farthest from the city center) and the remaining neighborhoods. Panel (a) groups them by region, while Panel (b) groups cities by development status. In both panels, each dot represents a city, and we plot group means along with 95 percent confidence intervals.

Panel (a) confirms the patterns we reported in Figure 3. The suburban–urban income gap is on average positive in the U.S. (around 15 percentile points), while it is on average zero in

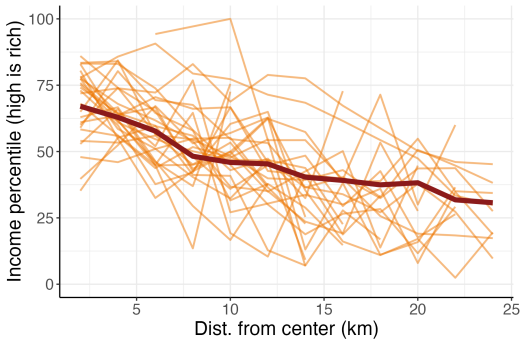
Figure 3: Residential Income and Distance to City Center, by Regions



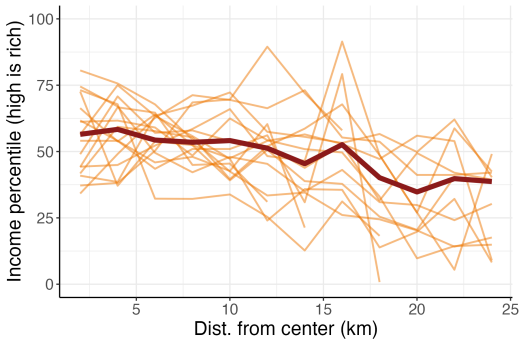
(a) U.S.



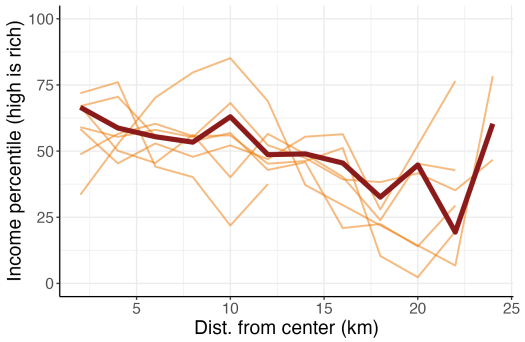
(b) W. Europe or Japan



(c) L. America



(d) Asia, E. Europe



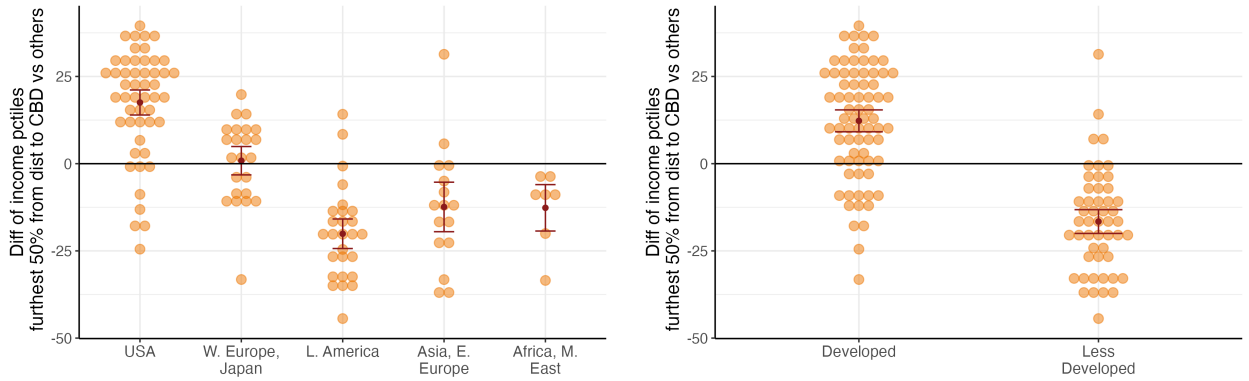
(e) Africa, M. East

Note: The figures show the relationships between distance from the city center and average neighborhood residential income percentile for each region, where panel (a) and (b) correspond to developed cities and (c)-(e) correspond to less-developed cities. Appendix Figure B.1 shows the same figures by aggregating developed and less-developed cities.

Western Europe and Japan.[8] Among less-developed regions, the gap is consistently negative across all regions. Panel (b) aggregates these patterns for developed and less-developed regions.

[8]In Appendix Figure B.7, we find that the U.S. pattern is similar between cities with above and below median shares of racial minorities (Black or Hispanic), suggesting that racial segregation is unlikely to be the main driver of this pattern.

Figure 4: Suburban-Urban Income Gap



(a) By Regions

(b) By Development Status

Note: The figures display the difference in average income percentiles between suburban and urban core neighborhoods for each city, where suburban areas are defined as the neighborhoods containing the 50 percent of the population located farthest from the city center, and urban core areas are defined as the rest. Each dot represents a city. Panel (a) groups them by region, while Panel (b) groups cities by development status. In both panels, we also report the group averages and their 95 percent confidence intervals. Appendix Figure B.4 shows the same set of figures using log income instead of income percentiles. Appendix Tables B.1 and B.2 report the values for each city and country, respectively.

In developed cities, the suburban–urban income gap is positive, averaging around 10 percentile points, while in less-developed cities, it is negative, at approximately minus 15 percentile points. Therefore, the differences between developed and less-developed cities are large and significant.

Appendix Tables B.1 and B.2 provide city- and country-level rankings of the suburban–urban income gap. Of the 20 cities with the most negative gaps, 18 are in less-developed countries; only Tokyo, Japan, and Seattle, U.S., are from developed countries. Conversely, 19 of the 20 cities with the largest positive gaps are in developed countries – all in the U.S. – with only one less-developed countries, Da Nang, Vietnam.

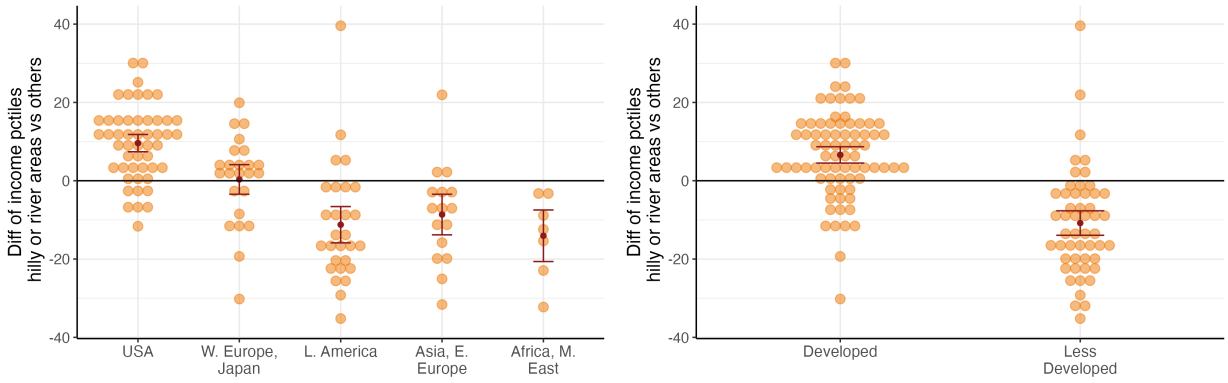
3.3. Residential Income and Hills/Rivers

We now turn to the relationship between residential income and natural geographic features, specifically hills and rivers. In the U.S., Lee and Lin (2018) show that such features are important predictors of neighborhood affluence: areas near hills and rivers tend to be wealthier than average. They interpret this pattern as reflecting the value households place on natural amenities, such as scenic views from elevated terrain or proximity to water. However, little is known about whether these patterns generalize to cities outside the U.S.

Figure 5 presents the difference in average income percentiles between neighborhoods located

in hilly or river-adjacent areas and those that are not, city by city. For simplicity, we combine the hilly and river neighborhoods in this figure, while the patterns are similar when looking at them separately (Appendix Figures B.5 and B.6). In the regression analysis that follows, we show that the patterns remain qualitatively similar when hills and rivers are analyzed separately.

Figure 5: Residential Income and Hills/Rivers: By City



(a) By Regions

(b) By Development Status

Note: The figures show the difference in average income percentiles between neighborhoods that are hilly or located near a river and those that are not. River neighborhoods are defined as areas within 100 meters of a natural waterway, while hilly neighborhoods are those with an average slope greater than 5 degrees. Each dot represents a city. Panel (a) groups them by region; Panel (b) groups cities by development status. In both panels, we report the group means along with 95 percent confidence intervals. Appendix Figures B.5 and B.6 show the same set of figures, separately for hills or rivers, respectively. Appendix Tables B.3 and B.4 report the values for each city and country, respectively.

Panel (a) shows that, on average, neighborhoods located in hilly or river-adjacent areas are approximately 10 percentile points richer than other neighborhoods in the U.S., consistent with the findings by Lee and Lin (2018). In Western Europe and Japan, these differences are zero on average, while there are large variations across cities. The differences between the U.S. and other developed cities may be driven by various factors, such as different types of amenities other than natural amenities emphasized in the U.S. context, or differences in commuting costs. While our main focus is on comparing developed and less-developed cities, in Section 6.4, we use our quantitative framework to examine the differences among developed cities.

In contrast to developed cities, the incomes in hilly or river neighborhoods are systematically lower than those in other neighborhoods, consistently across all three regions. By aggregating less-developed cities altogether in Panel (b), we find that the hilly and river areas exhibit lower income by 10 percentile points in less-developed cities, and they are statistically significantly different from developed cities.

Appendix Tables B.3 and B.4 report these gaps at the city and country levels, respectively. Of the 20 cities with the most negative income differences in hilly or river neighborhoods, 18 are located in less-developed countries, with only two from a developed country (Newcastle, U.K., and Toulouse, France). Conversely, among the 20 cities with the largest positive gaps, 18 are in developed countries, with Campinas, Brazil, and Da Nang, Vietnam, as the sole exceptions from a less-developed country.

3.4. Regression Results: Developed vs Less-Developed Cities

We now assess whether the contrasting pattern of spatial income distribution between developed and less-developed cities holds in a multiple regression framework that jointly accounts for all geographic features. Specifically, we estimate the following regression:

$$\text{Income}_{j,c} = \beta' X_{j,c} \times \text{Developed}_c + \gamma' X_{j,c} \times \text{Less Developed}_c + \nu_c + \epsilon_{j,c}, \tag{1}$$

where c indexes a city, j indexes neighborhoods, $\text{Income}_{j,c}$ denotes the proxy for residential neighborhood income (income percentiles and log residential income), and $X_{j,c}$ includes indicators for suburban, hilly, and river-adjacent areas, as defined in Section 2. The specification includes city fixed effects ν_c, and the error term $\epsilon_{j,c}$ captures idiosyncratic neighborhood-level variation. We weight observations by the inverse of the fraction of residents in each neighborhood within each city, such that the regression assigns equal weight to each city. We cluster standard errors at the city level.

Table 2 presents the results. Columns (1) use income percentile rank as the outcome, while Columns (2) use log average residential income. For each specification, the top panel reports the estimated coefficients from equation (1), and the bottom panel reports the differences in coefficients between developed and less-developed cities.

Column (1) shows that the associations between income and suburban, hilly, and river locations remain robust when these geographic features are jointly included in the regression. When disaggregated between hilly and river areas, the coefficients for developed cities are larger for hilly areas (13.7 percentile points) than for river areas (3.3 percentile points), with both estimates statistically significant. In contrast, for less-developed cities, the coefficients are similar and negative (-6.7 and -6.1 percentile points, respectively). In both cases, the differences between developed and less-developed cities are statistically significant and sizable, as reported in the bottom panel.

Column (2) shows that these patterns are robust when using log average income instead of income percentiles as the dependent variable. They also indicate economically meaningful magnitudes:

for example, Column (2) shows that suburban areas are associated with incomes that are 0.17 log points higher in developed cities, and 0.27 log points lower in less-developed cities. Overall, the results indicate robust and statistically significant differences in income premiums associated with suburban, hilly, and river neighborhoods between less developed and developed cities.

Table 2: Regression Results of Residential Income on Suburban, Hilly, and River Dummies

Dependent Variables:	Income percentile (high is rich)	Log average income
Model:	(1)	(2)
Variables		
$\text{Developed}_c \times \text{Suburban}_{j,c}$	12.6***	0.17***
	(1.9)	(0.03)
$\text{Less Developed}_c \times \text{Suburban}_{j,c}$	-15.4***	-0.27***
	(2.1)	(0.04)
$\text{Developed}_c \times \text{Hilly}_{j,c}$	13.7***	0.17***
	(3.3)	(0.04)
$\text{Less Developed}_c \times \text{Hilly}_{j,c}$	-6.7**	-0.11**
	(2.8)	(0.05)
$\text{Developed}_c \times \text{River}_{j,c}$	3.3***	0.06***
	(1.2)	(0.01)
$\text{Less Developed}_c \times \text{River}_{j,c}$	-6.1***	-0.09***
	(1.9)	(0.03)
Difference: Less Developed$_c$ vs Developed$_c$		
$\text{Suburban}_{j,c}$	-28.0***	-0.44***
	(2.8)	(0.05)
$\text{Hilly}_{j,c}$	-20.4***	-0.28***
	(4.3)	(0.07)
$\text{River}_{j,c}$	-9.5***	-0.16***
	(2.3)	(0.04)
Observations	49,649	49,649
Unique Cities	122	122
City FE	✓	✓
Weight by neighborhood pop within city	✓	✓

*Note: Top panel reports the results of the regression (1). Bottom panel reports the differences in the coefficients between developed and less-developed cities. The unit of observation is a neighborhood. We weight observations by the inverse of the fraction of residents in each neighborhood for each city, such that the regression assigns equal weight to each city. Standard errors are clustered at the city level. ***, ** and * indicate statistical significance at the 1-percent, 5-percent and 10-percent levels.*

Robustness In Table 3, we show that the differences between developed and less-developed cities remain robust across a variety of alternative specifications and samples. Row (1) replicates the baseline estimates from the bottom panel of Column (2) in Table 1. Our results are robust to restricting neighborhoods within 15 km and 20 km from the city center (Rows 2 and 3), indicating that the definition of city boundaries or differences in city size across development levels do not drive our results. Excluding U.S. and Brazil attenuates the differences between developed and less-developed cities (Rows 4 and 5), while the differences are statistically significant except for river coefficients when excluding U.S.

Table 3: Robustness: Differences in Income Premiums in Suburban, Hilly, and River Neighborhoods between Less Developed versus Developed Cities

		Difference: Less Developed vs. Developed		
	Specification	Suburban	Hilly	River
1	Baseline	-28.0 (2.8)***	-20.4 (4.3)***	-9.5 (2.3)***
2	Exclude neighborhoods $\geq$ 15km of center	-21.0 (2.8)***	-19.9 (5.4)***	-8.4 (2.6)***
3	Exclude neighborhoods $\geq$ 20km of center	-25.5 (2.8)***	-20.0 (4.9)***	-8.4 (2.5)***
4	Exclude USA	-16.1 (3.2)***	-13.8 (5.8)**	-5.1 (3.2)
5	Exclude Brazil	-24.4 (3.4)***	-16.3 (5.9)***	-11.4 (2.6)***
6	New World cities	-37.5 (3.4)***	-29.7 (4.5)***	-10.2 (3.1)***
7	Old World cities	-11.8 (3.8)***	-4.8 (6.5)	-7.0 (3.5)**
8	Control: neighborhood area x city FE	-25.0 (2.7)***	-21.1 (4.3)***	-6.9 (2.0)***
9	Control: neighborhood quadrant to center x city FE	-28.0 (2.8)***	-22.2 (4.4)***	-9.1 (2.1)***
10	Control: city population x neighborhood attributes	-27.0 (3.1)***	-20.2 (4.4)***	-9.3 (2.4)***
11	Control: city area x neighborhood attributes	-27.2 (2.9)***	-19.2 (4.5)***	-8.8 (2.4)***

*Note: This table presents robustness checks for the bottom panel of Column (2) in Table 1, which reports differences in the coefficients on suburban, hilly, and river dummies between less developed and developed cities, estimated using regression (1). Row (1) reproduces the baseline results from the bottom panel of Column (2) in Table 1. Rows (2) and (3) exclude neighborhoods located more than 15 km and 20 km from the city center, respectively. Rows (4) and (5) exclude cities in the U.S. and Brazil, respectively. Rows (6) and (7) restrict the sample to "New World" cities (North and Latin America) and "Old World" cities (all others), respectively. Rows (8) and (9) augment regression (1) by adding controls for neighborhood log area and for the quadrant relative to the city center (north, south, east, or west) interacted with city fixed effects, respectively. Rows (10) and (11) augment regression (1) by adding controls for log city population and geographic area, interacted with suburban, hilly, and river dummies, respectively. All regressions cluster standard errors at the city level. ***, ** and * indicate statistical significance at the 1-percent, 5-percent and 10-percent levels.*

Rows (6) and (7) restrict the sample to "New World" cities (North and Latin America) and "Old World" cities (all others), respectively, which has been noted as a key factor shaping subnational economic geography (Henderson, Squires, Storeygard, and Weil, 2018). While the patterns are

stronger among New World cities, they remain statistically significant in Old World cities as well, with the exception of hilly coefficients. Row (8) adds controls for neighborhood log area, interacted with city fixed effects, to address concerns that differences in neighborhood definitions or geographic boundaries might bias the results. Row (9) includes controls for quadrant location within the city (north, south, east, or west), interacted with city fixed effects, given prior evidence that neighborhood orientation influences income patterns (Heblich, Trew, and Zylberberg, 2021). In both cases, the main results are unaffected. Finally, Rows (10) and (11) augment regression (1) by including interactions between the suburban, hilly, and river dummies and log city population or geographic area. The results remain robust, suggesting that differences in city size, either in population or land area, do not explain the observed income patterns across development levels.

Additional Statistics Appendix B.4 investigates additional statistics related to income. Among less-developed cities with available household-level data, we find no significant association between suburban, hilly, or riverside locations and average household size or employment status (Appendix Tables B.5 and B.6). In contrast, these areas tend to have residents who are younger and less educated on average (Appendix Tables B.7 and B.8), suggesting that the observed income patterns may reflect sorting by age and skill. Population density in suburban, hilly, and riverside areas is lower in both developed and less-developed cities, though the rate of population decay differs: suburban areas exhibit faster declines, while hilly areas show slower declines in less-developed cities (Appendix Table B.9).

4. Commuting Frictions and Job Access

We next examine detailed bilateral commuting patterns across neighborhoods. We start by estimating the following "commuting gravity" equations for each city c separately:

$$\log \mathbb{E}\left[\lambda_{jn,c}\right] = \psi_{n,c} - \kappa_c \text{RoadDistance}_{jn,c} + \eta_{j,c} \tag{2}$$

Here, $\lambda_{jn,c}$ is the fraction of commuters who live in neighborhood j and work in neighborhood n in city c; $\text{RoadDistance}_{jn,c}$ is the road distance between j and n measured using OSRM as described in Section 2; $\psi_{n,c}$ and $\eta_{j,c}$ are origin and destination fixed effects, respectively; and $\mathbb{E}$ is the conditional expectation operator with respect to residuals. The coefficients on the road distance κ_c capture the semi-elasticity of commuting with road distance, i.e., the average percentage point decline in the probability of commuting associated with an additional one-kilometer increase in distance. In our theoretical model in Section 5, equation (2) is microfounded by households' optimal discrete commuting choice decisions, where $\eta_{n,c}$ is proportional to the

wage rates at workplace j, and $\psi_{n,c}$ is proportional to the negative of the expected wage rate conditional on residing in n. We leverage these structural interpretations for our quantitative analysis in Section 6.

We estimate this equation using a Pseudo-Poisson Maximum Likelihood (PPML) estimator to handle the presence of neighborhood pairs with zero commuters (Dingel and Tintelnot, 2020).[9] For this analysis, we exclude cities from Spain and Brazil due to a lack of bilateral commuting flow data, except for the city of Belem in Brazil, where the JICA travel survey is available.

4.1. Commuting Frictions

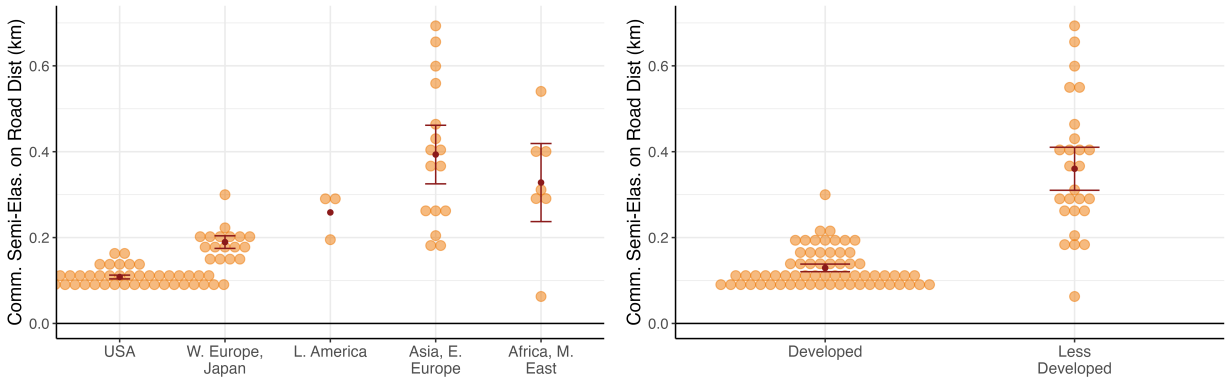
Figure 6 presents the estimated commuting semi-elasticities. As before, each dot represents an estimate from each city, where Panel (a) groups them by region, while Panel (b) groups them by development status. In Panel (a), we find stark differences in commuting semi-elasticities across regions. In the U.S., the semi-elasticity of commuting with respect to road distance is approximately 0.1, implying that an additional kilometer reduces commuting by 10 percent, holding workplace attractiveness constant. In Western Europe and Japan, the semi-elasticities are around 0.2, which is somewhat higher than in the U.S., potentially reflecting differences in reliance on private cars versus public transit. In developing countries, the average semi-elasticity rises to 0.35, with a large variation across cities. Despite this heterogeneity, the semi-elasticities in less-developed cities are significantly larger than those in developed cities, as shown in Panel (b).[10]

In Figure 7, we assess how the estimated commuting semi-elasticities relate to the traffic speed and commuting modes. In Panel (a), we plot the estimated commuting semi-elasticities against the "speed index" from Akbar et al. (2023b), which measures log-point differences in average road speed across cities using Google Maps API. We observe a tight negative relationship: in cities with faster average road speeds, commuting semi-elasticities tend to be lower. We find that the regression slope of the log of commuting semi-elasticities on speed index (in log scale) is -1.83. Interestingly, the slope is significantly larger than one in absolute terms, suggesting that

[9] In Appendix C.2, we assess how the sampling rates of travel surveys may affect the estimates of κ_c through a Monte Carlo simulation. We find that the statistical uncertainty associated with sample size is negligible. In Lima, the confidence interval of the estimates based on simulated data centers around the assumed true value (0.28 from our point estimate using actual data), with a confidence interval that covers less than 0.01 given the sample size (approximately 66,000).

[10] See Appendix Table C.1, C.2, C.3 of the point estimate for each city. The cross-region patterns are robust to using straight-line distance instead of road distance (Appendix Figure C.1), using log road distance instead of the level (Appendix Figure C.2), and estimating the coefficients on bins of distances nonparametrically instead of a single index (Appendix Figure C.3). Consistent with these findings, commuting distances are significantly longer for the developed cities than less-developed cities (Appendix C.4), although those of W. Europe and Japan are more similar to those of less-developed cities, likely because of more dispersed employment in the former cities.

Figure 6: Estimated Semi-Elasticity of Commuting to Road Distance (κ_c)



(a) By Regions

(b) By Development Status

Note: The figures show the estimated semi-elasticity of commuting to road distance using PPML estimator of specification (2). Each dot represents an estimate for each city. Panel (a) groups them by region, while Panel (b) groups them by development status. Sample cities exclude Spain and Brazil (except for the city of Belem) due to a lack of bilateral commuting flow data. In both panels, we report the group means along with 95 percent confidence intervals, except for "L. America" in Panel (b), where we only show the mean due to the small sample size.

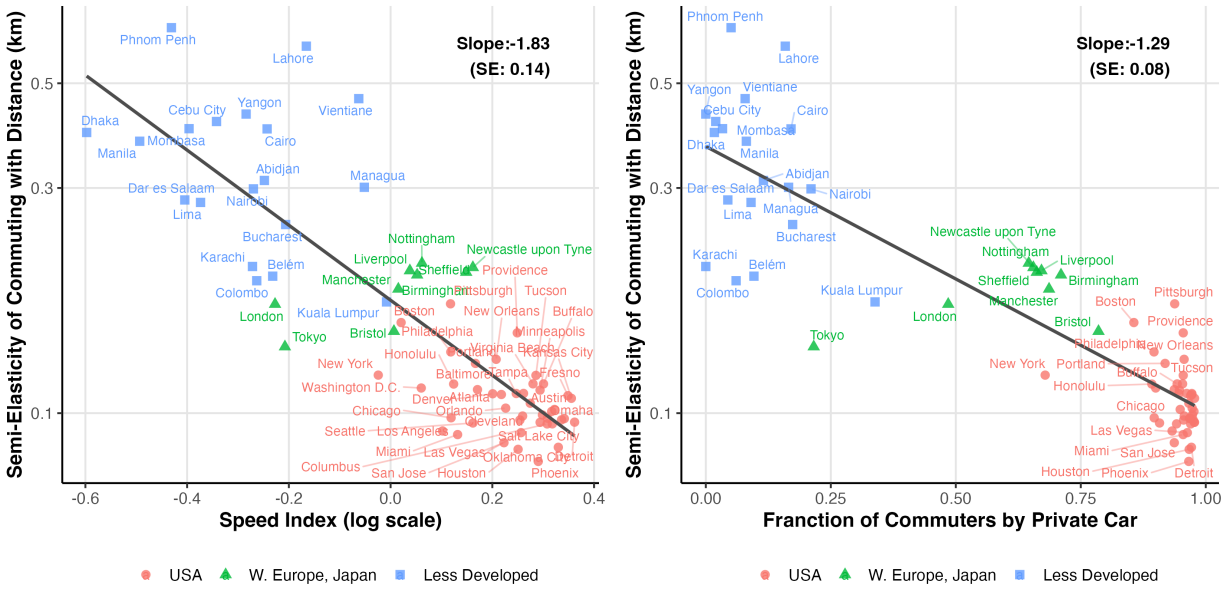
commuting frictions increase more than proportionally as road speeds decline. One potential explanation for this pattern is that, in faster cities, residents are more likely to rely on efficient transportation modes such as private vehicles, whereas in slower cities, individuals may be constrained to slower options like walking.

To further assess the role of transportation mode, in Panel (b), we plot the estimated commuting semi-elasticities against the fraction of commuters by private car. We measure this value using our travel surveys whenever they are available, and for other cities, we use the data constructed by Prieto-Curiel and Ospina (2024) based on various surveys and administrative data. We find a strong, significant negative relationship. Therefore, the choice and availability of cars appear to be another key determinant of the commuting semi-elasticity.

4.2. Job Access

The higher commuting semi-elasticity in less-developed cities suggests that neighborhoods differ markedly in job access within those cities. We now analyze how job access relates to residential income, continuing the analysis in Section 3. Specifically, we construct a measure of job access of neighborhood j in city c as the distance-discounted sum of workplace fixed effects (Tsivanidis,

Figure 7: Commuting Semi-Elasticity, Speed, and Modes



(a) Speed

(b) Commuting by Car

Note: The figures show the patterns of the estimated semi-commuting elasticities from specification (2). In Panel (a), we plot them against the "speed index" from Akbar et al. (2023b), which measures log-point differences in average road speed across cities using Google Maps API. In Panel (b), we plot them against the fraction of commuters by private car. We measure this value using our travel surveys whenever they are available, and for other cities, we use the data constructed by Prieto-Curiel and Ospina (2024) based on various surveys and administrative data. In both panels, we restrict to cities where both of these data are available. "Slope" and "SE" indicate the regression slopes of the log of semi-elasticity against each variable, as well as the standard errors for those estimates. Appendix Table C.4 reports the results based on joint regressions.

2025), given by

$$\text{JobAccess}_{j,c} = \log\left(\sum_{n} \exp(\hat{\kappa}_c \text{RoadDistance}_{jn,c} + \hat{\psi}_{n,c})\right), \tag{3}$$

where $\hat{\kappa}_c$ and $\hat{\eta}_{j,c}$ are estimates of κ_c and $\psi_{n,c}$ from equation (2). Under the PPML specification, the negative of the estimated origin fixed effects $\hat{\eta}_{j,c}$ are numerically equivalent (up to scale) to the geometric sum of the workplace fixed effects and the distance-related commuting costs: $\text{JobAccess}_{j,c} = -\hat{\eta}_{j,c}$ (Fally, 2015). Since the level of the fixed effects $(\hat{\eta}_{\ell,c}, \hat{\psi}_{j,c})$ are not identified from equation (2), we normalize these objects for each city to mean zero. As mentioned before, in our structural model in Section 5, this measure of job access is proportional to the log expected wage per efficiency unit of labor for residents in j.

Figure 8 presents the relationship between the residential income and the percentile point of job

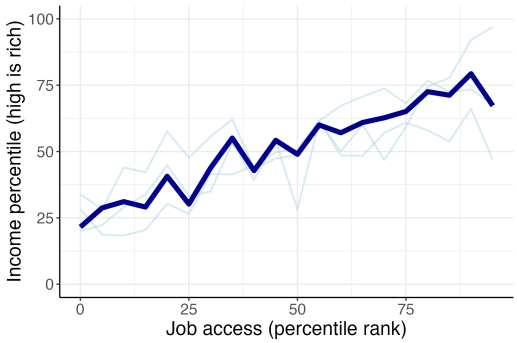
Figure 8: Residential Income and Job Access, by Region



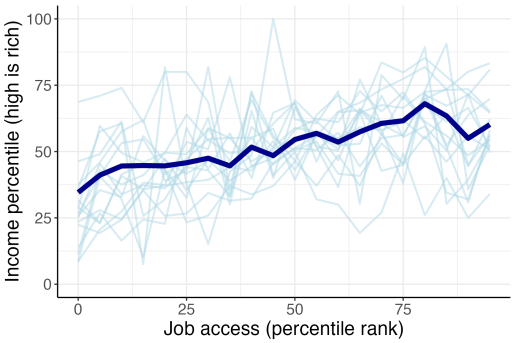
(a) USA



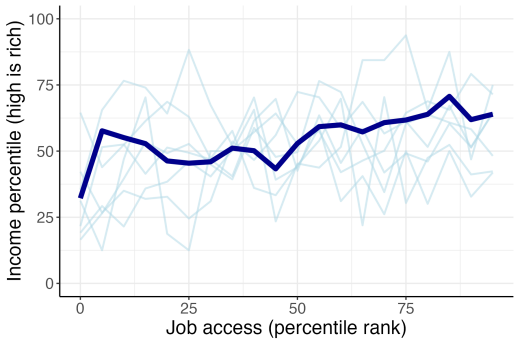
(b) W. Europe, Japan



(c) L. America



(d) Asia, E. Europe



(e) Africa, M. East

Note: The figures show the relationships between the percentiles of estimated job access (equation 3) within each city and average neighborhood residential income percentile for each region, where panel (a) and (b) correspond to developed cities and (c)-(e) correspond to less-developed cities. Each light line represents a single city, and averages are highlighted in bold. Appendix Figure C.6 presents the figures for developed and less-developed cities. Appendix Figure C.7 shows the income premiums in neighborhoods with above-median job access.

access within each city, separately for each region. We find that the income is monotonically decreasing in job access in developed cities, both in the U.S. (Panel a) and Western Europe and Japan (Panel b). In contrast, the opposite is true in less developed cities, consistently across

regions (Panels c-e). This mirrors our earlier finding that in less-developed cities, neighborhoods located away from the city center or near hills and rivers – often remote from employment hubs – tend to be poorer.[11] Interestingly, in contrast to the flat relationships of income with distance to city center, hills, and rivers for Western Europe and Japan, we find a downward-sloping relationship for job access, similar to the ones in the U.S.[12] These patterns indicate that, in those cities, there are rich neighborhoods that tend to be farther from employment but not necessarily far from city center or in hills and rivers – potentially reflecting other dimensions of amenities.

These results have important implications for the unequal benefits of urban space: in less-developed countries, poorer residents tend to reside in areas with worse access to jobs, while richer residents dominate in the city center, potentially exacerbating income inequality. In contrast, in developed countries, poorer residents often live closer to employment, and wealthier residents tend to move away from city center or neighborhoods with high amenities. We now turn to develop and estimate a quantitative spatial urban model to understand the drivers of these patterns and their welfare implications.

5. Model

In this section, we develop a quantitative theoretical framework to analyze the drivers of spatial income distribution and commuting patterns within the city.

Consider a city c that consists of $j \in \mathcal{J}_c$ neighborhoods. Each neighborhood j is endowed with exogenous amenity $B_{j,c}$, the productivity of final goods $A_{j,c}$, and the supply shifter of housing $S_{j,c}$. Furthermore, each pair of locations $j, n \in \mathcal{J}_c$ is endowed with commuting costs $\tau_{jn,c}$. For notational brevity, we omit subscripts c in this section and reintroduce them in the next section for the quantitative analysis.

Ex-ante heterogeneous households with different earning potentials choose residential locations based on job access, housing costs, and neighborhood amenities. They also supply labor to various locations, depending on the commuting costs and wages at the destinations, following a discrete choice formulation of commuting decisions (e.g., Ahlfeldt, Redding, Sturm, and Wolf, 2015; Tsivanidis, 2025). We further incorporate non-homothetic preferences over housing and amenities, implying that the relative importance of job access, housing, and amenities for

[11] In Appendix Table C.5, we show that the job access tends to be indeed lower in hills and rivers in both developed and less developed cities, while the penalty is larger in less-developed cities.

[12] Appendix Figure C.7 shows that the income premiums in above-median-job-access neighborhoods are negative and significant in "USA" and "W. Europe and Japan" independently, and they are positive and significant in less developed cities.

residential location choice shifts as household income rises.

The model's primary objective is to establish a mapping from the neighborhood fundamentals to the equilibrium spatial distribution of income and commuting patterns. In particular, we discuss alternative hypotheses about why the spatial income distribution differs between developed and less-developed cities. In Section 6, we take our model to data to quantitify each channel.[13]

5.1. Households' Preferences and Residential Location Decisions

There is a unit measure of households ω. Households are heterogeneous with respect to idiosyncratic preferences for residential location choice $\epsilon(\omega) \equiv \{\epsilon_j(\omega)\}_j$ and with respect to efficiency unit of labor $s(\omega)$, which we call "earning potential" for short. Each household decides sequentially where to reside, where and how much to supply labor, and how much to consume housing and freely traded final goods (numéraire).

We specify households' preferences over final goods and housing using non-homothetic constant elasticity of demand (NH-CES) preferences (Albouy, Ehrlich, and Liu, 2016; Comin, Lashkari, and Mestieri, 2021; Finlay and Williams, 2025; Hoelzlein, 2023). Non-homotheticity in housing is a robust empirical regularity and it has been pointed out as a force for gentrification and income sorting in the U.S. (Couture et al., 2024). Specifically, given the consumption amount of final goods y and housing h, the sub-utility of households U_j derived from the consumption of final goods and housing is implicitly determined by the following equation:

$$1 = \left(\frac{y}{U_j}\right)^{\frac{\sigma-1}{\sigma}} + \chi^{\frac{1}{\sigma}} \left(\frac{h}{U_j^{\varepsilon}}\right)^{\frac{\sigma-1}{\sigma}} \tag{4}$$

Here, σ determines the elasticity of substitution between housing and final goods. χ regulates the relative demand for housing. $\varepsilon(>0)$ is the parameter that governs the degree of non-homotheticity in housing. When $\varepsilon = 1$, equation (4) reduces to standard CES preferences. In the parameter range $0 < \sigma < 1$, where housing and final goods are complements (as documented by Finlay and Williams, 2025), housing is a subsistence good if and only if $0 < \varepsilon < 1$ (i.e., its expenditure share declines with income, holding prices fixed). The NH-CES preference class provides well-defined demand functions over any positive value of income and prices, and hence suitable for our applications that involve large dispersion of income and prices within and across

[13] In practice, the neighborhood-specific fundamentals are in part shaped by the endogenous supply of amenities (e.g., schools, restaurants), productivity spillovers (e.g., knowledge diffusion), and policy choices (e.g., infrastructure investment). The primary purpose of our model is not to model these processes explicitly, but rather to account for how each of these objects, inferred from our data, maps to the differences in observed spatial income distribution across cities (Section 6). For this purpose, it is not crucial to endogeneous these objects. Endogeneizing them becomes more important when addressing explicit policy counterfactuals.

cities.[14]

Given this sub-utility from consumption U_j, household ω's overall utility by residing in j is given by:

$$\log\left(U_j^{\rho} + B_j^{\rho}\right)^{\frac{1}{\rho}} + \epsilon_j(\omega) \tag{5}$$

where B_j is an exogenous amenity in location j, capturing factors such as the natural amenities available in hills or rivers or the open space available in suburban areas. ρ regulates the non-homotheticity of residential location choice with respect to B_j. To see why, if $\rho < 0$, $\frac{\partial}{\partial B_j}\frac{\partial}{\partial \log U_j}\log\left(U_j^{\rho} + B_j^{\rho}\right)^{\frac{1}{\rho}} > 0$, i.e., the elasticity of overall utility with respect to consumption subutility U_j increases in amenity B_j. Therefore, individuals with a higher earning potential, and hence U_j, tend to value an increase in amenity B_j more. If $\rho < 0$, the opposite is true. In the limit as $\rho \to 0$, the utility function converges to an additive form: $\log\left(U_j^{\rho} + B_j^{\rho}\right)^{\frac{1}{\rho}} \to \log U_j + \log B_j$, as in Tsivanidis (2025); Finlay and Williams (2025); Couture et al. (2024). We revisit how this property shapes the patterns of residential location choices in the next subsection.

We now describe households' decisions. First, conditional on residential locations, they decide how much to consume housing h and final goods y subject to the budget constraint:

$$\begin{aligned} U_j(s(\omega)) \equiv \arg\max_{\{h,y\}} \quad & U_j \\ \text{s.t.} \quad & P_j h + y \leq \overline{w}_j s(\omega) \quad \text{and equation (4)} \end{aligned} \tag{6}$$

where P_j is the housing rent; $s(\omega)$ is the earning potential of household ω; and $\overline{w}_j$ is the wage rate per efficiency unit of labor for residents in j, which is determined by the labor supply decision as we discuss in Section 5.3.

Anticipating this decision, household ω chooses the residential location that maximizes the utility (5)

$$j(\omega) \equiv \arg\max_j V_j(s(\omega)) + \epsilon_j(\omega), \qquad V_j(s) \equiv \log\left(U_j(s)^{\rho} + B_j^{\rho}\right)^{\frac{1}{\rho}} \tag{7}$$

Following the literature, we assume that $\epsilon_j(\omega)$ is independently drawn from the Gumbel distribution with scale parameter ν. Then, the probability that households with earning potential

[14] Alternative classes of non-homothetic preferences commonly used in studies of individual cities or countries, such as Stone-Geary preferences (Tsivanidis, 2025) or unit-demand preferences (Couture et al., 2024), are not well suited for our application. Utilities under these formulations are not well-defined over certain ranges of income and prices, and cannot accommodate large income differences across cities, as we consider here.

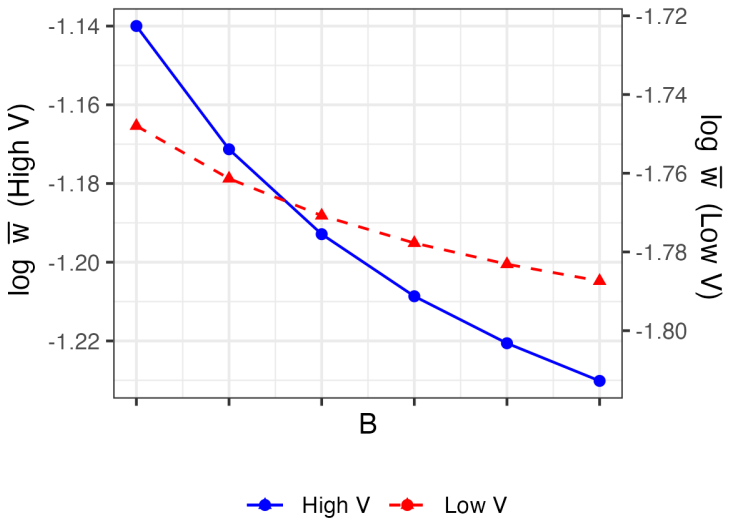
s reside in location j is given by

$$\pi_j(s) = \frac{\exp\left(\nu V_j(s)\right)}{\sum_\ell \exp\left(\nu V_\ell(s)\right)} \tag{8}$$

5.2. Non-homotheticity in Residential Location Choice

The residential location choice probability $\pi_j(s)$ depends on the individual's earning potential s because of the non-homothetic preferences with housing and amenity. To illustrate this point, in Figure 9, we plot the indifference curves between the wage rate $\overline{w}_j$ and amenity B_j. Specifically, assuming $\rho < 0$, the figure shows the combinations of $\{\overline{w}_j, B_j\}$ that deliver the same level of overall utility $V_j \in \{V^{High}, V^{Low}\}$, holding housing rents P_j fixed. Although the indifference curve for V^{High} lies above that for $V^{Low}(< V^{High})$, we normalize the vertical axis for each curve for ease of exposition.

Figure 9: Indifference Curves between $\overline{w}_j$ and B_j with $\rho < 0$



Note: The figures show the combinations of $\{\overline{w}_j, B_j\}$ that deliver the same level of utility (excluding idiosyncratic taste shocks) corresponding to $V_j \in V^{High}, V^{Low}$, where $V^{High} > V^{Low}$, holding housing rents P_j fixed, for the case with $\rho < 0$.

The figure demonstrates that the trade-off between wages $\overline{w}_j$ and amenities B_j varies with the overall utility level. When the overall utility is low, the indifference curve is relatively flat, indicating that agents are more responsive to wages than to amenities in their residential choice. Conversely, when utility is high, the indifference curve becomes steeper, suggesting a greater

sensitivity to amenities. This pattern reflects the non-homothetic nature of preferences: with $\rho < 0$, B_j behaves like a "luxury" good in location choice, valued more by higher-utility (or higher-income) individuals. When $\rho > 0$, the reverse holds.[15]

The "luxury" feature of amenity ($\rho < 0$) offers a potential explanation for why the spatial income distribution within cities varies depending on overall income levels. If the city's overall income level is sufficiently high, households with greater earning potential place higher value on amenity-rich locations, while those with lower earning potential sort into areas with amenity-scarce, higher-wage locations. In contrast, if the city's overall income level is sufficiently low, households place limited value on amenities and base their residential choices primarily on wages, regardless of their earning potentials.

Furthermore, in the equilibrium, housing rents P_j are endogenously determined by supply and demand in the housing market, and hence neighborhoods with higher residential demand tend to exhibit higher housing rents, as we further describe below. When housing is a subsistence good ($1 > \epsilon > 0$), as robustly documented in prior work, this creates an additional force toward gentrification. Households with lower earning potential s are more sensitive to housing costs due to their higher expenditure share on housing (Couture et al., 2024; Finlay and Williams, 2025).

5.3. Commuting (Labor Supply) Decisions

Each household ω consists of a continuum of members of unit measure, each endowed with $s(\omega)$ efficiency units of labor.[16] Members independently choose their work location. If member v decides to commute to work location $n = n(v)$, she earns income at wage rate $w_n \tilde{\epsilon}_n(v)$ per efficiency unit of labor, where $\tilde{\epsilon}_n(v)$ captures idiosyncratic productivity at that workplace. She also incurs commuting costs in the form of iceberg earnings losses, $\tau_{jn} \geq 1$, where j denotes the household's residential location. These costs reflect both distance and variation in transportation infrastructure – such as road quality in suburban, hilly, or river-adjacent areas. Together, the labor supply decision of a member v is given by

$$n(v) = \arg\max_n \tau_{jn}^{-1} w_n \tilde{\epsilon}_n(v). \tag{9}$$

[15] At first glance, it may seem puzzling that the constant-elasticity-of-substitution specification with B_j (equation 5) gives rise to non-homotheticity. This arises because B_j is fixed at the location level and does not require spending out of the household's budget, allowing individuals of any income level to access it conditional on residing there.

[16] We assume a continuum of household members, instead of a discrete number, to eliminate ex-post heterogeneity in wage rates conditional on residential location. Although this assumption can be relaxed without difficulty, it serves to simplify the exposition of residential location decisions in Section 5.1.

We assume that $\tilde{\epsilon}_n(v)$ is drawn from an i.i.d. Frechet distribution with shape parameter θ. Then, the probability that a household member residing in location j commuting to location n is given by

$$\lambda_{jn} = \frac{\left(\tau_{jn}^{-1} w_n\right)^{\theta}}{\sum_{\ell} \left(\tau_{j\ell}^{-1} w_{\ell}\right)^{\theta}}. \tag{10}$$

Notice that this equation provides a microfoundation for our empirical gravity equations in Section 4. Specifically, by parametrizing that $\tau_{jn,c}$ is a power function of road distance such that $\tau_{jn,c} = (\text{RoadDistance}_{jn,c})^{\tilde{\kappa}_c}$,

$$\log \lambda_{jn,c} = \theta \log w_{n,c} - \tilde{\kappa}_c \theta \text{RoadDistance}_{jn,c} - \log \sum_{\ell} \left(\tau_{j\ell,c}^{-1} w_{\ell,c}\right)^{\theta}, \tag{11}$$

which corresponds to equation (2), where $\psi_{n,c} \equiv \theta \log w_{n,c}$ is the workplace fixed effects; $\eta_{j,c} \equiv -\log \sum_{\ell} \left(\tau_{j\ell,c}^{-1} w_{\ell,c}\right)^{\theta}$ is the origin fixed effects; and $\kappa_c \equiv \tilde{\kappa}_c \theta$.

Furthermore, applying the law of large numbers, the wage rate per efficiency unit of labor for residents in j is given by

$$\log \overline{w}_j = \varrho + \log \sum_{\ell} \left(\tau_{j\ell}^{-1} w_{\ell}\right)^{\theta}, \tag{12}$$

where $\varrho \equiv \log \Gamma\left(\frac{\theta-1}{\theta}\right)$, where $\Gamma(\cdot)$ is the Gamma function. Therefore, the empirical job access measure defined in equation (3) coincides with the log of expected wage rate, up to scale.

5.4. Production, Market Clearing, and Equilibrium

Final goods are produced in each location j by perfectly competitive firms with linear production technology using labor with productivity A_j. Perfect competition implies that

$$w_j = A_j. \tag{13}$$

Housing is supplied by perfectly competitive developers using land, owned by the absentee landlord, and the final goods. Furthermore, the efficiency of housing supply S_j may vary across neighborhoods, reflecting differences in development costs driven by local geographic features, such as hills or proximity to rivers. We assume that the inverse supply function of housing is given by

$$P_j = \frac{1}{S_j} H_j^{\mu}, \tag{14}$$

where H_j is the aggregate supply of housing.

The market clearing of housing in location j is given by

$$H_j = \int_s h_j(s)\,\pi_j(s)dG(s), \tag{15}$$

where $G(\cdot)$ is the cumulative distribution function of earning potential $s(\omega)$ across households, and $h_j(s)$ is housing consumption by residents in j with earning potential s.

The equilibrium is defined by households' consumption $\{h_j(s), y_j(s)\}$, residential choice probabilities $\{\pi_j(s)\}$, labor supply probabilities $\{\lambda_{jn}\}$, wages $\{w_j\}$, and house prices $\{P_j\}$, that satisfy households' optimal consumption decision (6), residential location decision (7) and (8), labor supply decision (9) and (10), final goods producers' optimality condition (13), and housing supply and market clearing conditions (14) and (15).

5.5. Spatial Income Distribution

The average residential income in neighborhood j is given by

$$\text{Income}_j = \overline{w}_j \frac{\int_s s\pi_j(s)dG(s)}{\int_s \pi_j(s)dG(s)}, \tag{16}$$

Therefore, the equilibrium residential income in neighborhood j is affected by two components. First, it depends on the wage rates per efficiency unit of labor in residential location $\overline{w}_j$, determined by Equation (12). This term is higher if neighborhood j is surrounded by neighborhoods that offer higher wages, or equivalently, higher productivity $w_n = A_n$. Notice that the variation of $\overline{w}_j$ is lower for cities with lower commuting costs on average $\{\tau_{jn}\}$. In an extreme case, if $\tau_{jn} = 1$ for all j, n, then $\overline{w}_j$ does not vary across locations.

Second, the residential income is affected by the average earning potential of households residing in the neighborhood j, $\int_s s\pi_j(s)dG(s)/\int_s \pi_j(s)dG(s)$. This component is shaped by the non-homotheticity in residential location choice, together with the distribution of wages, housing costs, and amenities, as discussed in Section 5.2.

Our model offers a structural interpretation of the spatial distribution of residential income analyzed in Section 3. In particular, it offers four potential explanations for the contrasting spatial income distribution between developed and less-developed cities. The first is *non-homothetic preference over amenities*, where households with high earning potential sort into high-amenity locations. Second, *commuting costs tend to be higher in less developed countries*, which disproportionately penalize expected wages $\overline{w}_j$ in peripheral areas, including hilly and river areas. Third, *jobs may be more centralized in developing cities*, with disproportionately lower

A_j in peripheral areas. Finally, *residential amenity value falls faster away from the city center and in hills and rivers in less developed cities*, with relatively lower B_j in peripheral areas. In the following section, we assess the quantitative relevance of each channel.

6. Quantitative Analysis

In this section, we take our model to data to assess the driving forces behind the contrasting spatial income distribution between developed and less-developed cities. First, we calibrate the model to match commuting and residential income patterns in U.S. cities. Second, we identify differences in overall productivity, commuting costs, and the spatial distribution of productivity between the U.S. and less-developed cities, with the latter two disciplined by the commuting gravity equations estimated in Section 4. Third, starting from the U.S.-calibrated model, we counterfactually offset these three differences. We also use the calibrated model to analyze the differences between the U.S. and other developed cities, as well as the distributional welfare implications for the city's overall productivity improvement.

6.1. Calibrating the Model to U.S. Cities

We calibrate the structural parameters and fundamental variables for U.S. cities using both existing estimates from the literature and our own estimates based on the data, as summarized in Table 4.

We calibrate the elasticity of substitution between housing and final goods to $\sigma = 0.52$, and the elasticity of nonhomotheticity in housing demand to $\varepsilon = 0.36$, based on Finlay and Williams (2025), who estimate these parameters from the Panel Study of Income Dynamics (PSID) for U.S. cities. A value of $0 < \sigma < 1$ implies that housing and final goods are complements, while $0 < \varepsilon < 1$ indicates that housing behaves as a subsistence good, consistent with robust empirical patterns observed in the U.S. and many other countries.

We assume that the distribution of earning potential $G(\cdot)$ follows a bimodal structure, with two mass points at 1 and 3 (normalized), each occurring with probability 0.5. This specification generates a degree of income variance comparable to that observed in the U.S. economy, as reported by Heathcote, Perri, and Violante (2010).

We set the shape parameter of the idiosyncratic productivity shock at $\theta = 5$, consistent with median estimates reported in the literature for both developed and less-developed cities.[17] Given θ, we can back out the productivity at workplaces using the workplace fixed effects from the

[17]The estimates from the prior research range from 2.2 to 8.3 (e.g., Ahlfeldt et al., 2015; Kreindler and Miyauchi, 2023; Severen, 2023; Tsivanidis, 2025).

Table 4: Calibrated Parameters Targeting U.S. cities

Symbol	Value	Description	Source
σ	0.52	Elasticity of substitution for housing	Finlay and Williams (2025)
ε	0.36	Elasticity of nonhomotheticity in housing	Finlay and Williams (2025)
$G(\cdot)$	{3,1} wp 0.5	Earning potential distribution	Variance of income
θ	5	Dispersion of idiosyncratic productivity shock	Literature
$\{\kappa_c\}$	–	Semi-elasticity of commuting cost to road distance	Gravity equation
$\{A_{j,c}\}$	–	Productivity	Gravity equation
χ	0.06	Preference shifter for housing	GMM
ν	11.7	Residential location choice elasticity	GMM
ρ	-0.33	Elasticity of nonhomotheticity in amenity	GMM
β_0	8.32	Amenity level	GMM
β_1	0.82	Amenity for suburban areas	GMM
β_2	1.14	Amenity for hilly areas	GMM
β_3	0.26	Amenity for river areas	GMM
μ	0.50	Inverse housing supply elasticity	Literature
$\{S_{j,c}\}$	–	Housing supply shifter	Rents

Note: Calibrated parameters and their sources. See the main text for the GMM procedure.

gravity equations $A_{n,c} \propto \exp\left(\psi_{n,c}/\theta\right)$, city by city.[18] We also parametrize the commuting cost as a power function of road distance (equation 11), and recover city-specific commuting cost semi-elasticity by $\kappa_c = \hat{\hat{\kappa}}_c/\theta$, where $\hat{\hat{\kappa}}_c$ is as estimated in Section 4 by equation (2). Notice that we can also back out wage rates at residential location $\overline{w}_{j,c} \propto -\exp\left(\eta_{n,c}/\theta\right)$.

We calibrate the inverse housing supply elasticity $\mu = 0.5$ based on a typical value estimated in the literature (e.g., Saiz, 2010). We also set the housing supply shifter $S_{j,c}$ for each location to be consistent with the observed rents $P_{j,c}$.

We estimate the remaining preference parameters and amenities using our income, commuting, and rents data from the U.S. We begin by introducing additional parametric assumptions for the amenity term $B_{j,c}$, modeling it as a function of observable geographic features highlighted in Section 3:[19]

$$\log B_{j,c} = \beta_0 + \beta_1 \text{Suburban}_{j,c} + \beta_2 \text{Hills}_{j,c} + \beta_3 \text{Rivers}_{j,c}, \tag{17}$$

[18] This procedure only reveals the relative $\{A_{j,c}\}_j$ within the city, but not the levels. We set the common level shifter across U.S. cities to replicate the observed mean income of U.S. cities. We also subtract the log neighborhood area size from the estimated fixed effects $\psi_{n,c}$ to offset the mechanical effect that the destination fixed effects proportionally increase in neighborhood area size (Kreindler and Miyauchi, 2023).

[19] We take the approach of parameterizing amenity $B_{j,c}$, rather than inferring them directly from residential location choices, as often implemented in the literature of quantitative spatial models (e.g., Redding, 2023). This choice reflects a data limitation: we do not observe residential location choice for each value of earning potential s, but only the average income at each location.

where $\text{Suburban}_{j,c}$, $\text{Hills}_{j,c}$, and $\text{Rivers}_{j,c}$ indicate the dummies for the suburban, hilly, and river, as defined for our income regression equation (1), and $\{\beta_0, \beta_1, \beta_2, \beta_3\}$ are common parameters across all U.S. cities.[20]

We estimate the nonhomotheticity in amenity preferences ρ, the amenity parameters $\{\beta_0, \beta_1, \beta_2, \beta_3\}$, the elasticity of residential location choice ν, and the housing preference shifter χ, using a generalized method of moments (GMM) procedure, to replicate the key properties of spatial income distribution observed in the U.S. cities. Specifically, using observed median rents $\{P_{j,c}\}$ from ACS and the estimated wage rates at residential locations $\{\overline{w}_{j,c}\}$ from the gravity equations, we solve for the equilibrium residential location choice $\{\pi_{j,c}(s)\}$ for each earning potential s given a candidate parameter vector $\Theta \equiv \{\beta_0, \beta_1, \beta_2, \beta_3, \rho, \nu, \chi\}$. We then construct a set of moments $g_{j,c}(\Theta)$, defined as the difference between model-implied and empirically observed values for key location-specific statistics: (1) log average residential income in location j, as well as its interaction with suburban, hilly, and river dummies; (2) the interaction of log average residential income and estimated job access; (3) the interaction with (2) and hilly and river dummies[21]; (4) average housing expenditure share of residents in location j; and (5) residential location choice elasticity with respect to job access.[22] Finally, we choose the value that minimizes the GMM objective function:

$$\hat{\Theta} = \min_{\Theta} \left\{ \left(\frac{1}{N} \sum_{n,c} g_{j,c}(\Theta) \right)' \mathbf{W} \left(\frac{1}{N} \sum_{n,c} g_{j,c}(\Theta) \right) \right\}, \tag{18}$$

where $\mathbf{W}$ is the two-step optimal weighting matrix.

While the GMM procedure jointly determines all parameters using all moments, specific moments are particularly informative about certain parameters. The first set of moments is particularly informative about the value for $\{\beta_1, \beta_2, \beta_3\}$, which determines the differential sorting across s to ward suburban areas, hills, and river-adjacent neighborhoods. The second moments are informative about the value for ρ, since a more negative ρ implies that households with lower

[20] $\{\beta_0, \beta_1, \beta_2, \beta_3\}$ can be interpreted either as direct preferences toward suburban open space and natural amenities, or the preferences toward endogenous amenities that emerge in those places (e.g., schools). In our accounting analysis below, we first ask how much the observed differences in the overall productivity, commuting costs, and the spatial productivity distribution can explain the gap between U.S. cities and less-developed cities, fixing these parameters. The residuals reflect the differences in $\{\beta_0, \beta_1, \beta_2, \beta_3\}$ between developed and less-developed cities, arising either from differences in innate preferences (e.g., culture) or endogenous amenities (e.g., schooling, pollution, or residential infrastructure).

[21] To address concerns that the placement of roads may be endogenously correlated with unobserved residential amenities, we use job access estimated using straight-line distance, instead of road distance, in equation (2).

[22] In the model, it is defined by $\frac{\partial \ln \pi_j(s)}{\partial \ln \overline{w}_j} = \nu \frac{U_j^\rho}{U_j^\rho + B_j^\rho}(1 - \pi_j(s)) \frac{\partial \ln U_j}{\partial \ln \overline{w}_j}$. In the data, we target this elasticity to 2.2 (averaged across earning potential s), consistent with estimates from prior studies examining residential responses to transportation network expansions (Tsivanidis, 2025; Severen, 2023).

earning potential are disproportionately more responsive to job access. The third set of moments is informative about the value of β_0, as the level of amenities determines differential responses across different s to job access. The fourth moment is informative about the preference shifter for housing χ. The fifth moment is informative about the residential location choice elasticity ν.

Table 4 provides the point estimates, which we use for our quantitative analysis. Consistent with the moment choices, Appendix D.1 demonstrates that our model accurately replicates the relationship between observed residential income and suburban, hilly, river, and job access, capturing the patterns documented in Sections 3 and 4.

6.2. Unpacking the Gap between U.S. and Less-developed Cities

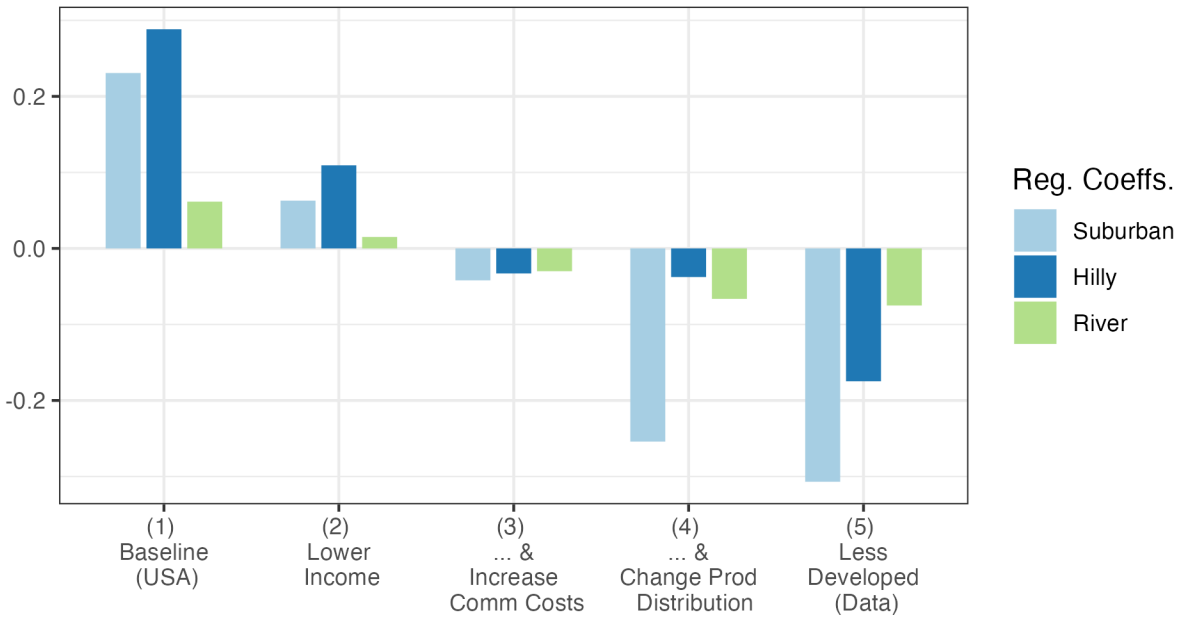
We now use our calibrated model for U.S. cities to conduct counterfactual simulations that eliminate differences in fundamentals between U.S. and less-developed cities. Specifically, we first recover $\{A_{j,c}\}$ and $\{\kappa_c\}$ for less-developed cities by applying the same procedure described above using the estimated commuting gravity equations. Next, we implement three sets of counterfactuals (as well as their combinations): (i) lowering the overall city productivity ($\{A_{j,c}\}$) to match the average income levels observed in less-developed cities; (ii) increasing commuting costs (κ_c) to the levels estimated for less-developed cities in Section 4; and (iii) adjusting the relative productivity penalty of neighborhoods in suburban, hilly, and river areas ($\{A_{j,c}\}$) to reflect the patterns observed in less-developed cities.

Figure 10 presents the results. In each column, we report the results of the estimated coefficients on suburban, hilly, and river dummies on residential income using the regression specification (1). Column (1) reports the patterns using our calibrated model to U.S. cities, which closely replicates the data pattern (Appendix Table D.1). Columns (2)-(4) report the regressions using U.S. cities under counterfactual equilibrium under alternative scenarios, as we further describe below. Column (5) reports the regression coefficients for less developed cities using our data, as reported in Table 2. Consistent with the findings so far, Column (1) exhibits positive income premiums in suburban, hills, and rivers for the U.S. cities, and Column (5) exhibits negative income premiums in those areas. We explore whether and what type of counterfactual closes this observed spatial income gap.

In Column (2), we present results from a counterfactual in which average productivity levels $\{A_{j,c}\}_j$ are uniformly reduced by 2.0 log points across all neighborhoods and cities – approximately 14% of the baseline values. This magnitude corresponds to the mean income gap between U.S. and less-developed cities in our sample.[23] Under this scenario, the regression

[23]Due to endogenous adjustments in residential choices, this counterfactual leads to an average income reduction of about 18% relative to the baseline, slightly greater than the assumed 14% decline in productivity.

Figure 10: Unpacking Spatial Income Distribution Gap: U.S. vs Less-developed Cities



Note: This figure displays the estimated coefficients on the suburban, hilly, and river indicators from regression specification (1). Column (1) presents the baseline results from our model calibrated to U.S. cities, corresponding to Column (1) of Table D.1. Column (2) shows the regression coefficients under a counterfactual in which average productivity levels $\{A_{j,c}\}_j$ are uniformly reduced by 2.0 log points across all neighborhoods and cities—approximately 14% of their baseline values. Column (3) reports coefficients from a counterfactual equilibrium in which we further increase the commuting semi-elasticity κ_c by 0.24, reflecting the average gap between U.S. and less-developed cities (Figure 6). Column (4) presents results from a counterfactual in which we further adjust the relative productivity of suburban, hilly, and river neighborhoods to match the patterns observed in less-developed cities (Table C.5, Column 2). Finally, Column (5) displays the corresponding estimates for less-developed cities based on observed data, as reported in Table 2.

coefficients on the suburban, hilly, and river indicators decline substantially and approach zero. This result is consistent with the interpretation that when overall income levels are low, even households with higher earning potential place less value on amenities, leading to weaker sorting into amenity-rich areas.

In Column (3), we run a counterfactual to additionally increase the commuting semi-elasticity κ_c by 0.24, matching the average difference in estimated κ_c between U.S. and less-developed cities (Figure 6). Under this scenario, the regression coefficients for suburban, hilly, and river areas become negative, consistent with the observation that higher commuting frictions in less

developed cities are contributing disproportionately worse job access in those areas.[24]

Column (4) reports results from a counterfactual in which we additionally adjust the relative productivity of suburban, hilly, and river neighborhoods to match the patterns observed in less-developed cities. While the estimated area-adjusted productivity $\{A_{j,c}\}$ is lower in suburban, hilly, and river areas, in both developed and less-developed cities, the latter cities feature disproportionately large penalties in suburban areas, to some extent, river and hilly areas (Column (2) of Table C.5). Consistent with these observations, we find that the negative income premiums in suburban areas become more pronounced with a more modest change in river and hilly areas. Together, these three forces account for 80 percent of the average differences in income premiums in suburban, hilly, and river neighborhoods between U.S. cities and less-developed cities.

While these three forces jointly explain much of the observed gap in spatial income distributions, some residual variation remains. These unexplained patterns likely reflect differences in amenity gradients within cities by development status. In contrast to developed cities, these areas may be perceived as less desirable due to environmental and infrastructural shortcomings—such as polluted rivers or the absence of sewage systems in suburban or hilly neighborhoods (Harari, 2024; McCulloch et al., 2025).[25]

To quantify the potential role of amenities, we ask: How much must the coefficients $\{\beta_1, \beta_2, \beta_3\}$ decline to account for the gap between Columns (4) and (5)? Table 5 presents the results. While these coefficients lack natural units and should be interpreted cautiously, the findings suggest that differences in amenity valuations may indeed contribute to the residual spatial income gaps. Notably, we do not need to assume negative amenity valuations for suburbs, hills, or rivers; rather, much of the observed gap is already explained by differences in income levels and commuting costs.

6.3. Heterogeneity of Commuting Costs and Wages by Income Groups

In our baseline analysis, we abstracted from income-related heterogeneity in commuting costs and wage distributions. However, in U.S. cities, disparities in transportation modes and commuting costs across income groups have been identified as one potential driver of residential

[24] Appendix Figure E.1 presents the results of a counterfactual exercise in which we implement the scenarios in Columns (2)–(4) individually, rather than cumulatively. We find that all of these counterfactuals tend to reduce income premiums, while the counterfactual to lower income has a larger effect than increasing commuting costs and changing productivity distribution.

[25] Another part of the gap between Columns (4) and (5) may arise from differences in the geographic structure of cities beyond commuting cost disparities between U.S. and less-developed cities; for example, the specific location of rivers and hills within the cities.

Table 5: Estimated $\{\beta_1, \beta_2, \beta_3\}$ to Rationalize the Income Distribution in Less-Developed Cities

Model	Suburban (β_1)	Hilly (β_2)	River (β_3)
Baseline Estimates (USA)	0.82	1.14	0.26
Estimates to Fully Rationalize Less Developed Cities	0.70	0.80	0.24

Note: First row reports the estimated $\{\beta_1, \beta_2, \beta_3\}$ using U.S. cities as reported in Table 4. Second row reports the estimated $\{\beta_1, \beta_2, \beta_3\}$ to fully rationalize the differences in the income premiums of suburban, hilly, and river neighborhoods, after accounting for the overall productivity differences, commuting cost differences, and differences in productivity premiums in those areas, as described further in Section 6.2.

sorting by income, where poorer residents without private cars sort into downtown areas (Glaeser et al., 2008; Su, 2022). If this tendency is weaker in less-developed cities, it could potentially explain the observed gap in spatial income distribution.

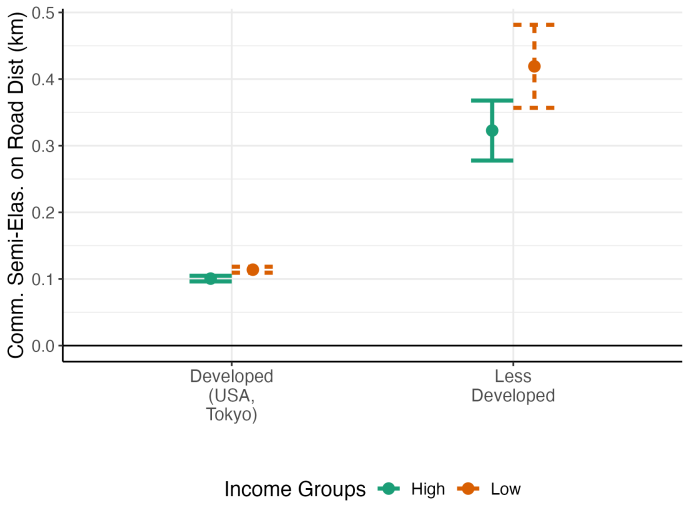
To assess this possibility, we estimate the empirical commuting gravity equation (2) separately for households with above- and below-median incomes within each city. Figure 11 reports the estimated commuting semi-elasticities for developed cities (U.S. cities and Tokyo; commuting data by income are not available for the U.K. and France) and for less-developed cities. Consistent with previous studies, we find larger commuting semi-elasticities among low-income groups (0.11) than high-income groups (0.10) in developed cities. However, this difference is *more* pronounced in less-developed cities, with 0.42 for low-income groups than 0.32 for high-income groups. These findings suggest that, while these forces may contribute to the sorting of low-income households toward central areas in both developed and less-developed cities, they are unlikely to account for the observed disparity between the two groups of cities.

To further assess these points, in Appendix F, we extend our model from Section 5 to allow commuting costs and wages to depend on individuals' earning potential s. Specifically, we now let commuting costs $\tau_{jn,c}(s)$ and wages $w_{n,c}(s)$ vary with s. This model predicts a separate commuting gravity equation for each s. We then calibrate the extended model using the commuting gravity equations by above-/below-median income for high/low earning potential groups.[26] Finally, we run the same counterfactual simulations as before, except that we change the commuting costs and spatial productivity distribution for each s.

Consistent with the observations above, we find that changing commuting costs and productivity distributions to the levels of less-developed cities generates slightly smaller effects on the spatial

[26]Realized income in our model is a product of earning potential s and the expected wage rates at the residential location $\overline{w}_{j,c}(s)$. Therefore, if the variance of earning potential s is much larger than the spatial variation of wage rates $\overline{w}_{j,c}(s)$, this strategy effectively splits the samples with high- and low-earning-potential households within each city.

Figure 11: Semi-Elasticity of Commuting by Income Groups



Note: This figure presents the estimated commuting semi-elasticities based on equation (2), separately for households with above and below median income for each city. The left column shows the mean estimates and 95% confidence intervals separately for high- and low-income households – defined as above or below the city-specific median income – in developed cities (U.S. cities and Tokyo; we do not have commuting data for U.K. and France separated by income), and the right column shows those estimates in less-developed cities.

income distribution compared to our baseline scenario (Figure 10). Therefore, the heterogeneity across earning potential s in commuting costs and spatial productivity distribution does not explain the differences in spatial income distribution. At the same time, our main conclusion remains: differences in aggregate income levels, commuting costs, and spatial productivity distributions *across cities* jointly account for the main observed gap in spatial income distributions between U.S. and less-developed cities.

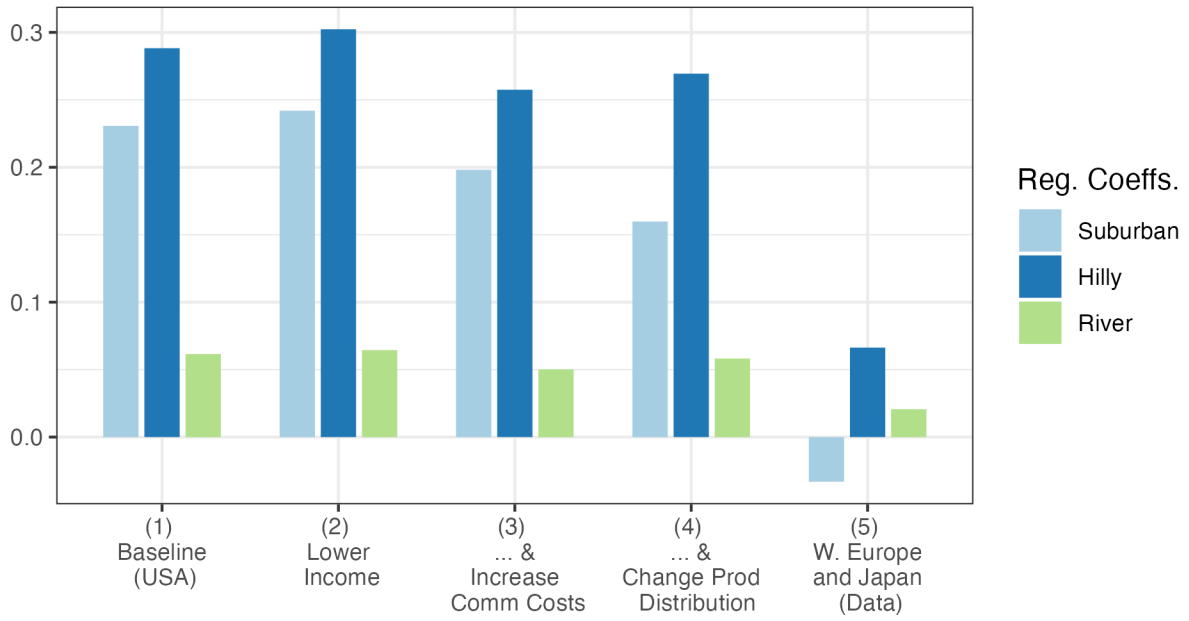
6.4. Unpacking the Gap between U.S. and Other Developed Cities

So far, we have focused on understanding the spatial income distribution patterns in less-developed cities, using U.S. cities as a benchmark for developed cities. At the same time, as we documented in Section 3 and 4, there is some heterogeneity between U.S. and non-U.S. developed cities. While the main focus of this paper is the comparisons between developed and less-developed cities, our framework can also be used to analyze the differences among developed cities. We now conduct this exercise.

In Figure 12, we undertake the same set of counterfactual exercises as before, except that we feed in the differences in overall income, commuting cost, and productivity distribution using

the estimates from Western Europe and Japan, instead of less-developed cities. In the latter cities, the regression coefficients are around zero and insignificant (Column 5).

Figure 12: Unpacking Spatial Income Distribution Gap: U.S. vs Other Developed Cities



Note: This figure displays the results of the same counterfactual simulation as in Figure 10, except that the counterfactuals are targeted to non-U.S. developed cities (Western Europe and Japan).

When we lower income to the level of other developed cities (a 0.19 log-point reduction) in Column (2), all regression coefficients increase slightly, in contrast to the counterfactual reduction to the level of less-developed cities (a 2.0 log-point decline). This non-monotonic response reflects the interplay between non-homothetic preferences for amenities and housing: lower income weakens the sorting force with respect to amenities but strengthens it with respect to housing costs. For small income reductions, the latter effect dominates. In Column (3), when we raise the commuting-cost semi-elasticity to the level of non-U.S. developed cities (an increase of 0.07), the coefficients decline meaningfully, though they remain well above zero. Finally, in Column (4), adjusting relative productivity in suburban, hilly, and river areas to match non-U.S. developed cities further reduces the suburban coefficient, consistent with more dispersed employment in those cities, though the coefficients remain far from zero.

Our quantitative analysis indicates that differences in the spatial income distribution between

these cities and the U.S. are primarily driven by the spatial distribution of amenities.[27] This suggests that, beyond the suburban open space and natural amenities emphasized in the U.S., other forms of amenities may play a central role elsewhere. For instance, cities like Paris or Amsterdam have a wide variety of restaurants, cafes, and bars in their urban cores, attracting wealthier residents (Brueckner et al., 1999; Almagro and Domínguez-Iino, 2025). This naturally raises the question of why the distribution of amenities systematically diverges between U.S. and other developed cities. One possibility is that commuting frictions shape endogenous amenity supply. Using smartphone data from Tokyo, Miyauchi, Nakajima, and Redding (2025) shows that commuters often combine workplace trips with visits to amenities, encouraging their agglomeration in central areas. Another explanation lies in the decentralization of local public goods policy: compared to the U.S., provision in many other developed countries is more centralized at the metropolitan level (OECD, 2021). Pinpointing the precise mechanisms behind these differences remains an important task for future research.

6.5. Distributional Impacts of Development

So far, we have focused on explaining the observed gap in spatial income distributions across cities. In this section, we show that these patterns also shape the unequal welfare gains associated with overall city development.

Specifically, using our model calibrated to U.S. cities, we compute welfare gains from a uniform increase in productivity – from the levels observed in less developed cities to the U.S. level, corresponding to a 2.0 log-point rise – as analyzed in Section 6.2. We examine how this aggregate productivity improvement translates into welfare gains across households, disaggregated by earning potential and residential location. We measure these gains using an equivalent variation (EV) metric that asks: "Given the existing U.S. distribution of population and housing prices, how much income would a household be willing to give up to avoid a citywide drop in productivity to the level of low-income cities, assuming no change in their residential location choice?" Following Baqaee and Burstein (2023), it is composed of two additive terms: (i) (uniform) changes in labor income across locations and household types, and (ii) location- and type-specific changes in expenditure-adjusted housing costs.

Our model predicts that this welfare measure varies across households due to nonhomothetic preferences, which induce shifts in spatial population distribution and generate uneven changes in housing costs. In a special case where we shut down nonhomotheticity (i.e., $\epsilon = 1$ and $\rho \to 0$), the welfare effects are uniform for all households, because wages and housing costs change

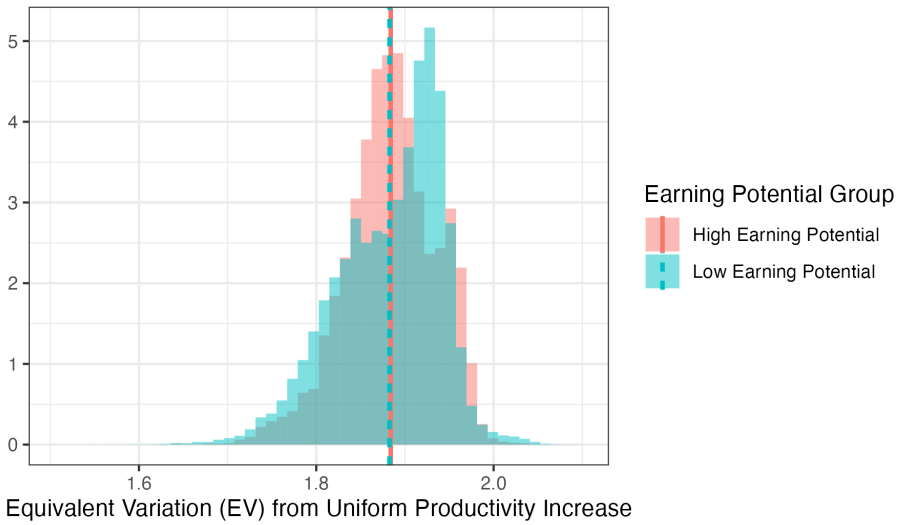
[27] Appendix Table E.1 presents an analysis analogous to Table 5, where we estimate $\{\beta_1, \beta_2, \beta_3\}$ to rationalize the patterns observed in other developed cities. The estimated parameters are substantially smaller, though still positive.

uniformly across locations.

Figure 13 shows the distribution of equivalent variation (EV), in log points, across households by residential location and earning potential. The average EV is approximately 1.9 log points, both for high-earning and low-earning potential groups. This value is somewhat below the assumed 2.0 log point increase in overall productivity. This gap reflects the fact that rising housing costs partially offset the benefits of higher productivity, abstracting from income gains to landowners.

The finding of similar average welfare gains across groups reflects the interplay of two opposing forces: (i) low-earning-potential households devote a larger share of income to housing, which tends to lower their EV, and (ii) they are more likely to reside in central urban areas with better job access in high-income (U.S.) cities, where housing costs rise less in response to a uniform productivity increase due to sorting. Our results suggest that the two forces roughly dominate.

Figure 13: Equivalent Variations (EV) from Uniform Productivity Increase



Note: The distribution of EVs, in log points, across households by residential location and earning potential, to increase the overall productivity from low-income-city level to high-income-city level (as observed in the U.S. cities).

We also find substantial dispersion in EVs across residential locations within each earning potential group. This reflects the spatial variation in housing cost changes induced by non-homothetic preferences. As the overall productivity rises, the housing costs in suburban, hilly, and river areas increase. This implies that lower-earning-potential households who live in those areas tend to have lower welfare gains.[28]

[28]Table E.2 reports the regression results of the EVs with the neighborhood characteristics for each earning potential group.

Taken together, these findings highlight that the evolution of spatial income distribution during economic development has important implications for the distribution of welfare gains across households.

7. Conclusion

How do internal city structures differ between developed and less-developed cities? Drawing on new granular data from 50,000 neighborhoods in 122 cities across 27 developed and less-developed countries. We construct this dataset using travel surveys from many developed-country cities, supplemented by compiling various sources of administrative data.

We document that, in less-developed cities, poorer residents are more likely to live farther from city centers and in hills and rivers – those with natural amenities but are nonetheless distant from jobs. In developed cities, these patterns are absent or even reversed. Commuting shares also fall more steeply with distance in less-developed countries, reflecting greater commuting frictions that exacerbate inequality in job access.

To interpret these findings, we develop a quantitative urban model that incorporates residential and commuting choices with non-homothetic preferences for housing and amenities. We find that the differences in urban spatial income distribution are largely driven by non-homothetic preferences over amenities, higher commuting costs, and more spatially concentrated jobs in less-developed cities. Our model also predicts that as cities grow richer, welfare gains are distributed unequally across households.

Our paper is among the first to provide cross-country comparisons of city structure and income distribution, spanning countries with low to high incomes. We hope it is only the beginning. In particular, we see two key directions for future research. First, while we document systematic differences in productivity, amenities, and commuting costs as important drivers of city structure and income distribution, we do not specify *why* these fundamentals differ between developed and less-developed cities. In practice, neighborhood-level fundamentals are partly shaped by the endogenous provision of amenities (e.g., schools, restaurants), productivity spillovers (e.g., knowledge diffusion), and policy choices (e.g., infrastructure investment). A more precise understanding of the sources of these differences is an important task for future work. Second, although our analysis is cross-sectional, our findings also inform the study of how the spatial organization of economic activity evolves with development. Historical evidence from developed countries (e.g., Lee and Lin, 2018) and theories of dynamic structural transformation of cities offer promising avenues for further exploration.

References

Adukia, A., S. Asher, K. Jha, P. Novosad, and B. Tan (2022): "Residential Segregation and Unequal Access to Local Public Services in India: Evidence from 1.5 m Neighborhoods," .

Ahlfeldt, G. M., T. Albers, and K. Behrens (2020): "Prime locations," *CEPR Discussion Paper No. DP15470*.

Ahlfeldt, G. M., N. Baum-Snow, and R. Jedwab (2023): "The skyscraper revolution: Global economic development and land savings," *CESifo Working Paper*.

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015): "The economics of density: Evidence from the Berlin Wall," *Econometrica*, 83, 2127–2189.

Akbar, P., V. Couture, G. Duranton, and A. Storeygard (2023a): "Mobility and congestion in urban India," *American Economic Review*, 113, 1083–1111.

Akbar, P. A., V. Couture, G. Duranton, and A. Storeygard (2023b): "The fast, the slow, and the congested: Urban transportation in rich and poor countries," *Working paper, National Bureau of Economic Research*.

Albouy, D., G. Ehrlich, and Y. Liu (2016): "Housing demand, cost-of-living inequality, and the affordability crisis," *Working Paper, National Bureau of Economic Research*.

Almagro, M. and T. Domínguez-Iino (2025): "Location sorting and endogenous amenities: Evidence from amsterdam," *Econometrica*, 93, 1031–1071.

Alonso, W. (1964): *Location and Land Use: Toward a General Theory of Land Rent*, Cambridge, MA: Harvard University Press.

Balboni, C., G. Bryan, M. Morten, C. O'Connor, and B. Siddiqi (2025): "Impacts of Infrastructure: People versus Place," *Working Paper*.

Baqaee, D. R. and A. Burstein (2023): "Welfare and output with income effects and taste shocks," *The Quarterly Journal of Economics*, 138, 769–834.

Baum-Snow, N. (2020): "Urban Transport Expansions and Changes in the Spatial Structure of U.S. Cities: Implications for Productivity and Welfare," *Review of Economics and Statistics*, 102, 929–945.

Becker, G. S. (1965): "A Theory of the Allocation of Time," *The economic journal*, 75, 493–517.

Bohr, C. E., M. Mestieri, and F. Robert-Nicoud (2024): "Economic Geography and Structural Change," *Working Paper*.

Bordeu, O. (2023): "Commuting infrastructure in fragmented cities," *Working Paper*.

Bright, E., A. Rose, and M. Urban (2016): *LandScan Global 2015*, Oak Ridge, TN: Oak Ridge National Laboratory, 2015 ed.

Brueckner, J. K., J.-F. Thisse, and Y. Zenou (1999): "Why is central Paris rich and downtown Detroit poor?: An amenity-based theory," *European economic review*, 43, 91–107.

Bryan, G., K. Frye, and M. Morten (2025): "Spatial Economics for Low- and Middle-Income Countries," Unpublished Working Paper, Stanford University.

Bryan, G., E. Glaeser, and N. Tsivanidis (2020): "Cities in the Developing World," *Annual Review of Economics*, 12, 273–297.

Busso, M., P. E. Carrillo, and J. P. Chauvin (2023): *Rethinking Urban Migration: Policy Options for Cities in Latin America and the Caribbean*, Latin American and Caribbean Microeconomic Report, Inter-American Development Bank.

Chatterjee, S., E. Giannone, and K. Kuno (2021): "Unequal global convergence," *Working Paper*.

Chauvin, J. P., E. L. Glaeser, Y. Ma, and K. Tobio (2017): "What is Different about Urbanization in Rich and Poor Countries? Cities in Brazil, China, India and the United States," *Journal of Urban Economics*, 98, 17–49.

Coeurdacier, N., F. Oswald, and M. Teignier (2025): "Structural change, land use and urban expansion," *Review of Economic Studies*.

Comin, D., D. Lashkari, and M. Mestieri (2021): "Structural change with long-run income and price effects," *Econometrica*, 89, 311–374.

Couture, V., C. Gaubert, J. Handbury, and E. Hurst (2024): "Income growth and the distributional effects of urban spatial sorting," *Review of Economic Studies*, 91, 858–898.

Davis, D. R. and J. I. Dingel (2020): "The Comparative Advantage of Cities," *Journal of International Economics*, 123.

Dingel, J. I., A. Miscio, and D. R. Davis (2021): "Cities, Lights, and Skills in Developing Economies," *Journal of Urban Economics*, 125, 103174.

Dingel, J. I. and F. Tintelnot (2020): "Spatial economics for granular settings," Tech. rep., National Bureau of Economic Research.

Eckert, F. and M. Peters (2022): "Spatial structural change," *Working Paper, National Bureau of Economic Research*.

Fally, T. (2015): "Structural gravity and fixed effects," *Journal of international economics*, 97, 76–85.

Finlay, J. and T. C. Williams (2025): "Housing demand, inequality, and spatial sorting," *Journal of International Economics*, 104171.

Florczyk, A., M. Melchiorri, C. Corbane, M. Schiavina, M. Maffenini, M. Pesaresi, P. Politis, S. Sabo, S. Freire, D. Ehrlich, T. Kemper, P. Tommasi, D. Airaghi, and L. Zanchetta (2019): "Description of the GHS Urban Centre Database 2015," *Publications Office of the European Union*.

Franklin, S., C. Imbert, G. Abebe, and C. Mejia-Mantilla (2024): "Urban Public Works in Spatial Equilibrium: Experimental Evidence from Ethiopia," *American Economic Review*, 114, 1382–1414.

Gechter, M. and N. Tsivanidis (2023): "Spatial spillovers from high-rise developments: Evidence from the Mumbai Mills," *Unpublished manuscript*.

Gertler, P., M. Gonzalez-Navarro, R. Undurraga, and J. A. Urrego (2025): "In-situ Upgrading or Population Relocation? Direct Impacts and Spatial Spillovers of Slum Renewal Policies," *Working Paper*.

Glaeser, E. L., M. E. Kahn, and J. Rappaport (2008): "Why do the poor live in cities? The role of public transportation," *Journal of urban Economics*, 63, 1–24.

Harari, M. (2024): "Residential Patterns and Local Public Goods in Urban Brazil," *Working Paper*.

Harari, M. and M. Wong (2024): "Slum Upgrading and Long-run Urban Development: Evidence from Indonesia," *Working paper*.

Heathcote, J., F. Perri, and G. L. Violante (2010): "Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006," *Review of Economic dynamics*, 13, 15–51.

Heblich, S., A. Trew, and Y. Zylberberg (2021): "East-Side Story: Historical Pollution and Persistent Neighborhood Sorting," *Journal of Political Economy*, 129.

Henderson, J. V., T. Squires, A. Storeygard, and D. Weil (2018): "The Global Distribution of Economic Activity: Nature, History, and the Role of Trade," *Quarterly Journal of Economics*, 133, 357–406.

Henderson, J. V. and M. A. Turner (2020): "Urbanization in the Developing World: Too Early or Too Slow?" *Journal of Economic Perspectives*, 34, 150–173.

Hoelzlein, M. (2023): "Two-sided sorting and spatial inequality in cities," *Working Paper*.

Jedwab, R., P. Loungani, and A. Yezer (2021): "Comparing cities in developed and developing countries: Population, land area, building height and crowding," *Regional Science and Urban Economics*, 86, 103609.

Khanna, G., A. Nyshadham, D. Ramos-Menchelli, J. A. Tamayo, and A. Tiew (2023): "Spatial mobility, economic opportunity, and crime," *Working Paper*.

Kreindler, G. E. and Y. Miyauchi (2023): "Measuring commuting and economic activity inside cities with cell phone records," *Review of Economics and Statistics*, 105, 899–909.

Larrick, G., Y. Tian, U. Rogers, H. Acosta, and F. Shen (2020): "Interactive Visualization of 3D Terrain Data Stored in the Cloud," in *2020 11th IEEE Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)*, 0063–0070.

Lebrand, M. and T. Kleineberg (2024): "The Distribution of Economic Activity across Cities: A Global Perspective," *Working Paper*.

Lee, S. and J. Lin (2018): "Natural Amenities, Neighbourhood Dynamics, and Persistence in the Spatial Distribution of Income," *Review of Economic Studies*, 85, 663–694.

Lehner, B. and G. Grill (2013): "Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems," *Hydrological Processes*, 27, 2171–2186.

Luxen, D. and C. Vetter (2011): "Real-time routing with OpenStreetMap data," in *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, New York, NY, USA: ACM, GIS '11, 513–516.

Margo, R. A. (1992): "Explaining the postwar suburbanization of population in the United States: The role of income," *Journal of urban economics*, 31, 301–310.

Marx, B., T. Stoker, and T. Suri (2013): "The Economics of Slums in the Developing World," *Journal of Economic Perspectives*, 27, 187–210.

McCulloch, S. E., M. P. Schaelling, M. Turner, and T. Kitagawa (2025): "Sewers and Urbanization in the Developing World," *Working Paper, National Bureau of Economic Research*.

Michaels, G., D. Nigmatulina, F. Rauch, T. Regan, N. Baruah, and A. Dahlstrand (2021): "Planning ahead for better neighborhoods: Long-run evidence from tanzania," *Journal of Political Economy*, 129, 2112–2156.

Miyauchi, Y., K. Nakajima, and S. J. Redding (2025): "The economics of spatial mobility: Theory and evidence using smartphone data," *Quarterly Journal of Economics*.

OECD (2021): *Fiscal Federalism 2022: Making Decentralisation Work*, Paris: OECD Publishing.

Prieto-Curiel, R. and J. P. Ospina (2024): "The ABC of mobility," *Environment International*, 185, 108541.

Redding, S. J. (2023): "Quantitative Urban Models: From Theory to Data," *Journal of Economic Perspectives*, 37, 75–98.

Rosenthal-Kay, J. (2024): "Urban costs around the world," Working Paper.

Saiz, A. (2010): "The geographic determinants of housing supply," *The Quarterly Journal of Economics*, 125, 1253–1296.

Severen, C. (2023): "Commuting, labor, and housing market effects of mass transportation: Welfare and identification," *Review of Economics and Statistics*, 105, 1073–1091.

Spear, B. D. (2011): "NCHRP 08-36, Task 098 Improving Employment Data for Transportation Planning," *American Association of State Highway and Transportation Officials*.

Su, Y. (2022): "The Rising Value of Time and the Origin of Urban Gentrification," *American Economic Journal: Economic Policy*, 14, 402–39.

Tabuchi, T. (2019): "Do the rich and poor colocate in large cities?" *Journal of Urban Economics*, 113, 103186.

Tokyo Metropolitan Area Transportation Planning Council (2018): "Tokyo Metropolitan Area Transportation Survey," Tech. rep.

Tsivanidis, N. (2025): "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogota's TransMilenio," Working Paper.

United Nations (2022): *World Cities Report 2022: Envisaging the Future of Cities*, World Cities Report, Nairobi, Kenya: United Nations.

Zárate, R. D. (2024): "Spatial Misallocation, Informality, and Transit Improvements: Evidence from Mexico City," *Working paper*.

Online Appendix for

"The Spatial Distribution of Income in Cities: New Global Evidence and Theory"

October 14, 2025

Peter Deffebach, David Lagakos, Yuhei Miyauchi, Eiji Yamada

Contents

A Data Appendix **1**

A.1 Additional Figures and Tables for Data 1

A.2 Comparison of Income between Travel Survey and Census in Belem, Brazil 6

A.3 Commuting Flow Data in the U.S.: CTPP vs LODES vs Travel Surveys 8

B Appendix for Spatial Distribution of Income **11**

B.1 Residential Income and Distance to City Center 11

B.2 Residential Income and Hills/Rivers 15

B.3 Separating U.S. by Racial Minority Shares 18

B.4 Other Statistics 19

C Appendix for Commuting Frictions and Job Access **22**

C.1 Commuting Frictions 22

C.2 Small Sample Properties of Commuting Semi-elasticity Estimates 28

C.3 Job Access 29

D Model Fit **31**

E Additional Figures and Tables for Quantitative Analysis **33**

F Heterogeneity of Commuting Costs and Spatial Productivity Distribution by Income Groups **35**

A. Data Appendix

A.1. Additional Figures and Tables for Data

Table A.1: List of Cities From JICA Surveys (L. America)

City Name	Year	Country	Number of respondents	Number of survey zones	Total geographic area (km^2)	Maximum distance to city center (km)
Belem	2000	Brazil	29835	91	417.0	21.8
Lima	2003	Peru	144490	388	1998.5	40.6
Managua	1998	Nicaragua	37082	92	244.7	13.8

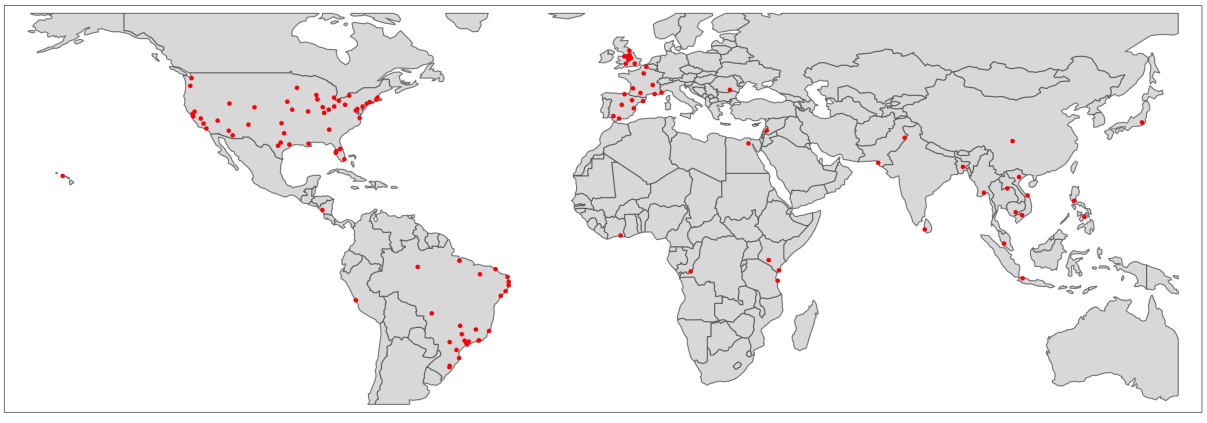
Table A.2: List of Cities From JICA Surveys (Africa, M. East)

City Name	Year	Country	Number of respondents	Number of survey zones	Total geographic area (km^2)	Maximum distance to city center (km)
Abidjan	2013	Côte d'Ivoire	50619	129	860.1	25.7
Cairo	2001	Egypt	137513	429	1576.3	36.5
Damascus	1998	Syrian Arab Republic	38280	74	1348.1	22.8
Dar es Salaam	2007	U.R. of Tanzania: Mainland	26687	159	1152.5	29.0
Kinshasa	2018	D.R. of the Congo	42031	321	538.8	21.4
Mombasa	2015	Kenya	10868	32	188.0	10.5
Nairobi	2013	Kenya	16794	102	590.7	20.1

Table A.3: List of Cities From JICA Surveys (Asia, E. Europe)

City Name	Year	Country	Number of respondents	Number of survey zones	Total geographic area (km^2)	Maximum distance to city center (km)
Bucharest	1998	Romania	92784	75	566.4	13.1
Cebu	2014	Philippines	28806	229	447.8	38.7
Chengdu	2000	China	31130	125	589.9	17.0
Colombo	2013	Sri Lanka	124673	376	1681.9	59.4
Da Nang	2008	Viet Nam	18171	50	368.3	14.3
Dhaka	2014	Bangladesh	118026	140	1536.6	29.7
Hanoi	2005	Viet Nam	63716	250	979.7	31.9
Ho Chi Minh	2014	Viet Nam	42908	210	1801.6	35.5
Jakarta	2018	Indonesia	5000	534	1939.0	65.5
Karachi	2011	Pakistan	96507	170	1369.4	30.3
Kuala Lumpur	1999	Malaysia	80545	222	2537.7	46.9
Lahore	2010	Pakistan	89414	188	1873.4	43.1
Manila	1996	Philippines	231838	220	1553.7	37.4
Phnom Penh	2012	Cambodia	42074	85	511.3	17.5
Viang Chan	2007	Lao People's DR	27630	33	297.6	11.0
Yangon	2013	Myanmar	42224	620	712.4	29.7

Figure A.1: Locations of Cities in our Data Set



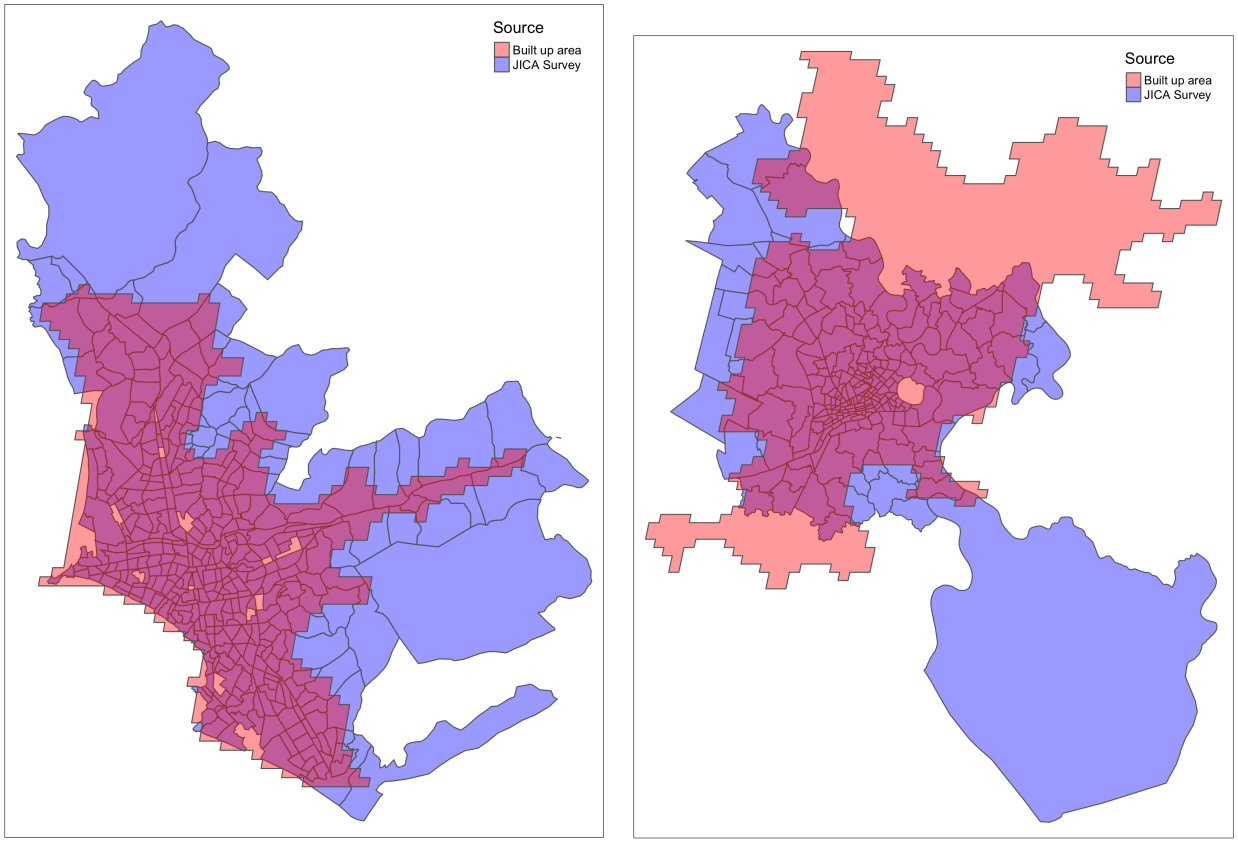
Note: A map of all cities included in the neighborhood-level income dataset described in Section 2 .

Table A.4: List of Cities Other Than JICA Surveys

Country	Number of Cities	Total number of survey zones	List of Cities
Brazil	24	3575	Aracaju, Belo Horizonte, Campinas, Cuiabá, Curitiba, Florianópolis, Fortaleza, João Pessoa, Londrina, Maceió, Manaus, Natal, Novo Hamburgo, Porto Alegre, Recife, Ribeirao Preto, Rio de Janeiro, Santos, Sao Goncalo, Sao Jose dos Campos, São Paulo, Teresina, Uberlândia, Vila Velha
France	7	562	Bordeaux, Lille, Lyon, Marseille, Nice, Paris, Toulouse
Japan	1	190	Tokyo
Spain	7	9889	Barcelona, Bilbao, Madrid, Málaga, Seville, Valencia, Zaragoza
United Kingdom	9	2529	Birmingham, Bristol, Leeds, Liverpool, London, Manchester, Newcastle upon Tyne, Nottingham, Sheffield
United States	48	27579	Albuquerque, Atlanta, Austin, Bakersfield, Baltimore, Boston, Bradenton, Bridgeport, Buffalo, Chicago, Cincinnati, Cleveland, Columbus, Concord, Dallas, Denver, Detroit, Fresno, Honolulu, Houston, Indianapolis, Kansas City, Las Vegas, Los Angeles, Louisville, Miami, Milwaukee, Minneapolis, New Orleans, New York, Oklahoma City, Omaha, Orlando, Philadelphia, Phoenix, Pittsburgh, Portland, Providence, Sacramento, Salt Lake City, San Antonio, San Jose, Seattle, St. Louis, Tampa, Tucson, Virginia Beach, Washington D.C.

Note: A list of all countries which appear in our neighborhood-level income dataset described in Section 2, along with the number of cities from each country.

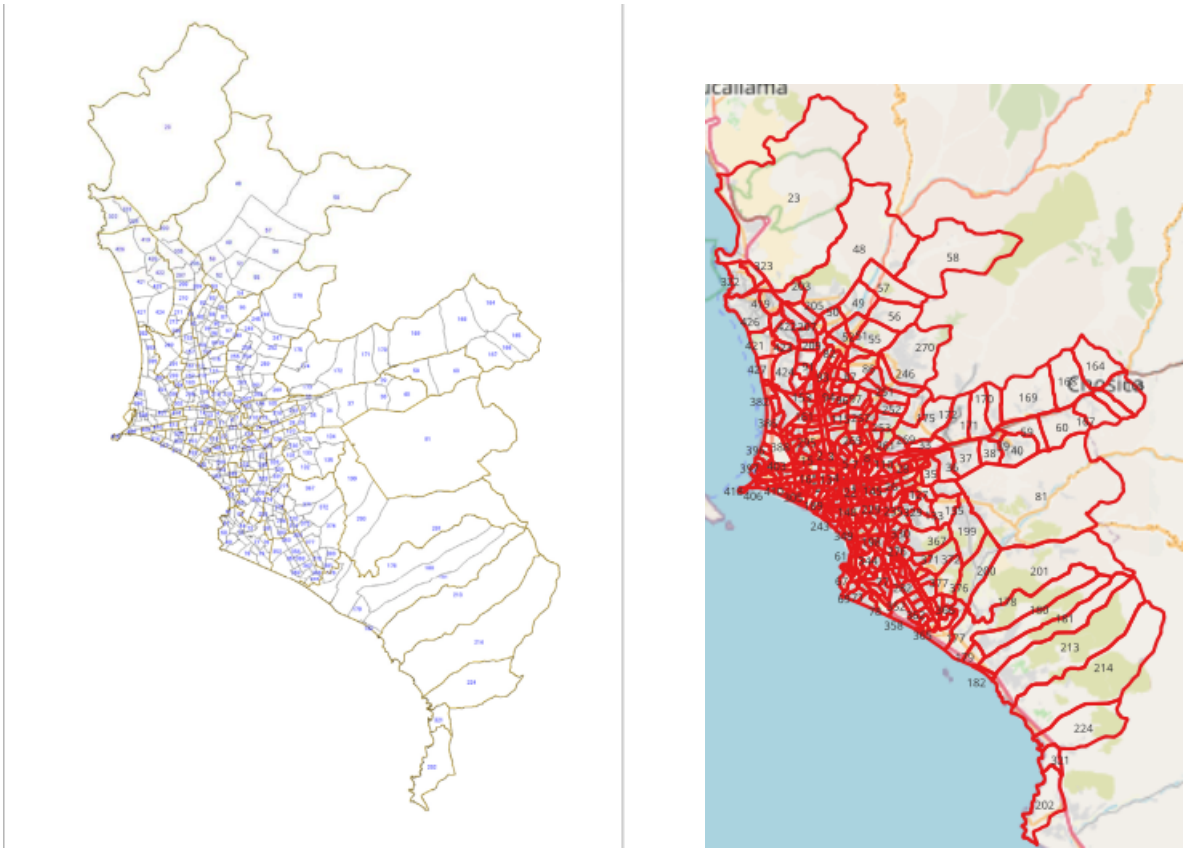
Figure A.2: Built-up area (red) and Neighborhoods (blue) in Lima and Ho Chi Minh City



(a) Lima, Peru (b) Ho Chi Minh City, Vietnam

Note: "Built-up Area", in red, represents shape of a city as defined by the Un Human Settlements Project (Bright et al., 2016). "JICA Survey", blue, shows the neighborhoods surveyed by JICA in their surveys across the world. With a few, small, deviations, the area spanned by JICA surveys encompasses the entire "built-up area".

Figure A.3: Geocoding Survey Zones (Lima)



(a) PDF (b) Shapefiles

Note: The Left panel shows the screenshot of survey zones from a report for a travel survey. The right panel shows the screenshot of the geocoded shapefiles of survey zones.

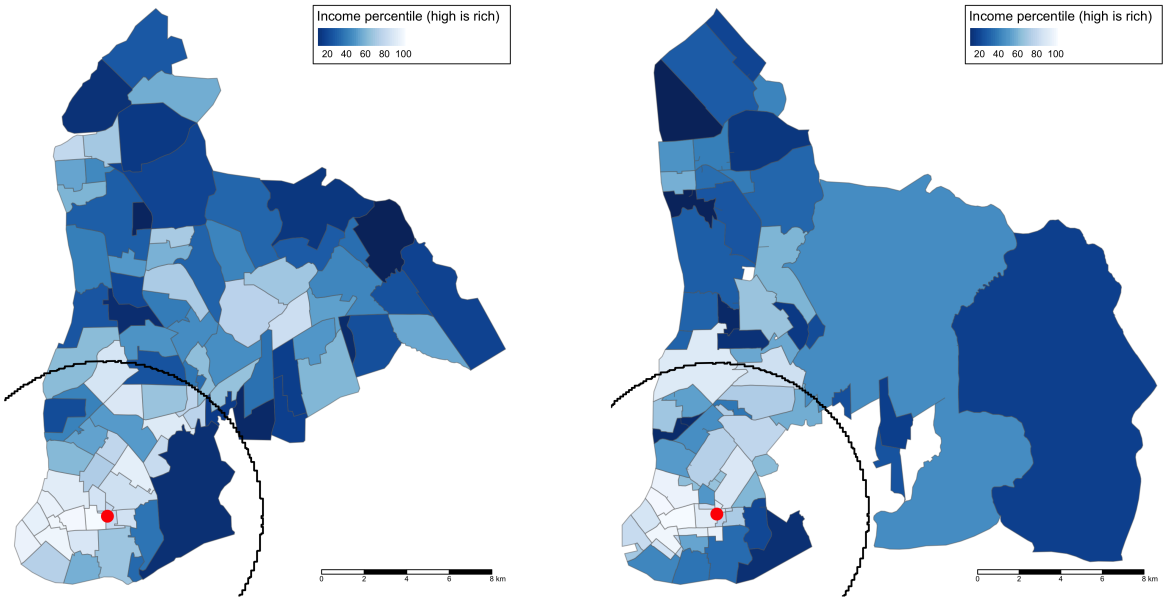
A.2. Comparison of Income between Travel Survey and Census in Belem, Brazil

In this section, we provide a cross-validation of our residential income data from travel survey data and census data in Belem, Brazil, the only city in our data set where both types of data is available.

The two data sets show similar patterns of residential income. In Figure A.4, we show the neighborhood-level income percentiles for both data sources. The semi-circle in each panel represents the border between suburban and non-suburban neighborhoods as defined in Section 3, where suburban neighborhoods are defined as those containing 50 percent of the population furthest from the city center. Despite the differences in the spatial resolution, one can recognize a similarity in the broad pattern of the spatial income distribution between these two data sets. In particular, in both data sets, one can visually recognize the pattern that the average income is higher in urban than in suburban areas.

To further reinforce this point, in Figure A.5, we show the relationship between income percentile and distance to the city center for the two data sources. The two thick lines represent the average income percentile within each 2 km bin from each data set, with 95 percent confidence intervals. Both datasets exhibit a monotonically decreasing pattern in the distance from the city center, consistent with our findings in a typical less-developed city (Section 3).

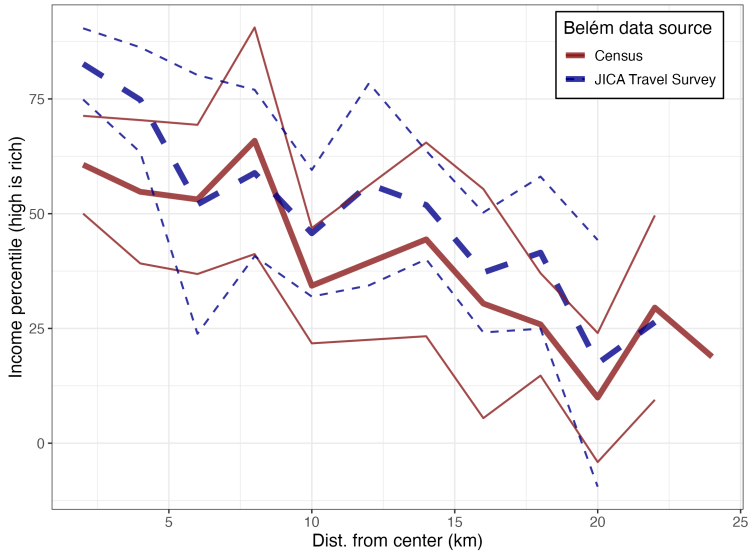
Figure A.4: Map of income percentiles in Belem, Brazil



(a) JICA Travel Survey

(b) 2010 Census

Figure A.5: Income percentile and distance from center in Belem



A.3. Commuting Flow Data in the U.S.: CTPP vs LODES vs Travel Surveys

For our analysis, we use commuting flow information in the U.S. from the Census Transportation Planning Packages (CTPP) for the years 2012-2016, which are constructed using the American Community Survey (ACS) data on usual residence and workplaces. The CTPP data reports the aggregate number of workers living and working in any given pair of census tracks.

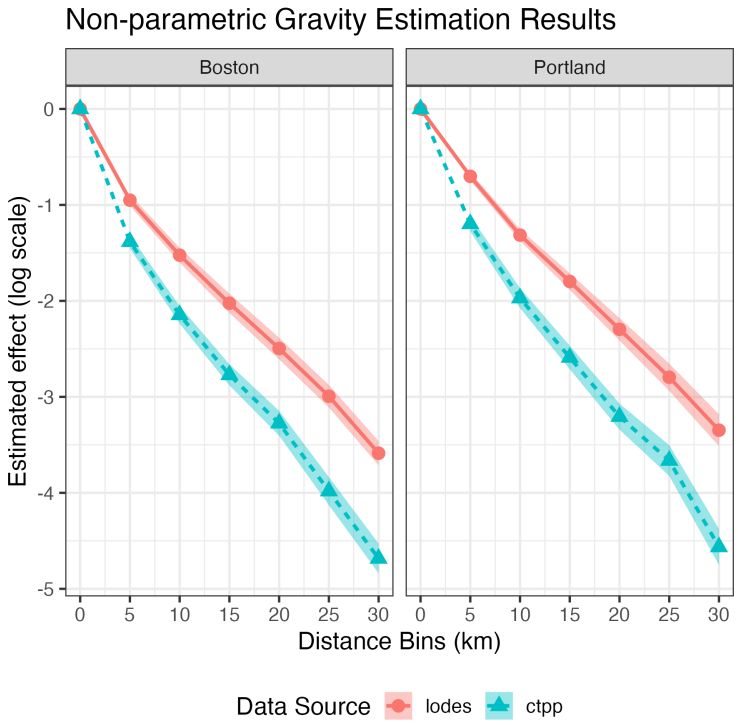
An alternative commuting dataset in the U.S. is the Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics (LODES). This dataset is constructed from the Quarterly Census of Employment and Wages, which covers all firms with paid employees and are subject to unemployment insurance laws. While LODES provides a near-census coverage of paid workers with their precise residential locations, it excludes the self-employed and imputes worksite locations for multi-establishment firms. Spear (2011) reports that, reflecting these differences, LODES tend to report longer commuting distances than CTPP. Below, we further compare these two data sets, focusing on the commuting semi-elasticities, and compare them with alternative travel survey data from two cities: Chicago and San Francisco.

In Figure A.7, we report the decay of commuting shares over road distances between the two commuting data sets for the cities of Boston and Portland. Specifically, we run a version of commuting gravity (2) using PPML estimator, where we replace the linear road distance in the dependent variable with dummies of distance bins of 5-10, 10-15, 15-20, 20-25, 25-30, and over 30 kilometers, where we exclude 0-5 kilometer as an omitted baseline category. We use 2015 data for LODES, similar in timing to our CTPP data (2012-2016). We find that, indeed, the LODES exhibit significantly shallower decay of commuting probability over road distance. In Figure A.7, we compare the commuting semi-elasticity estimates following the specification of A.7 using two different data sets. We find that the commuting semi-elasticities using LODES are roughly two-thirds the magnitude of those estimated from the CTPP.

Finally, in Figure A.8, we further compare the commuting semi-elasticities from these two data sets with the ones estimated from travel surveys for two cities, Chicago (conducted in 2007) and San Francisco (conducted in 2000), where that data is available.[29] Note that the timing of these surveys does not align with LODES and CTPP, and hence the estimates do not have to perfectly align with each other. Nonetheless, we find that the semi-elasticities from travel surveys align more closely with the CTPP data, with estimates from LODES are systematically lower than the other two data sets.

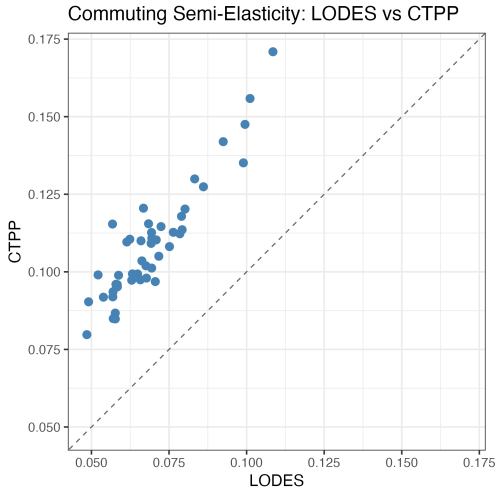
[29]The data is available from https://www.nrel.gov/transportation/secure-transportation-data/tsdc-chicago-household-travel-inventory and https://www.icpsr.umich.edu/web/ICPSR/studies/34805.

Figure A.6: Nonparametric Gravity Regression of Commuting Flows: CTPP vs LODES



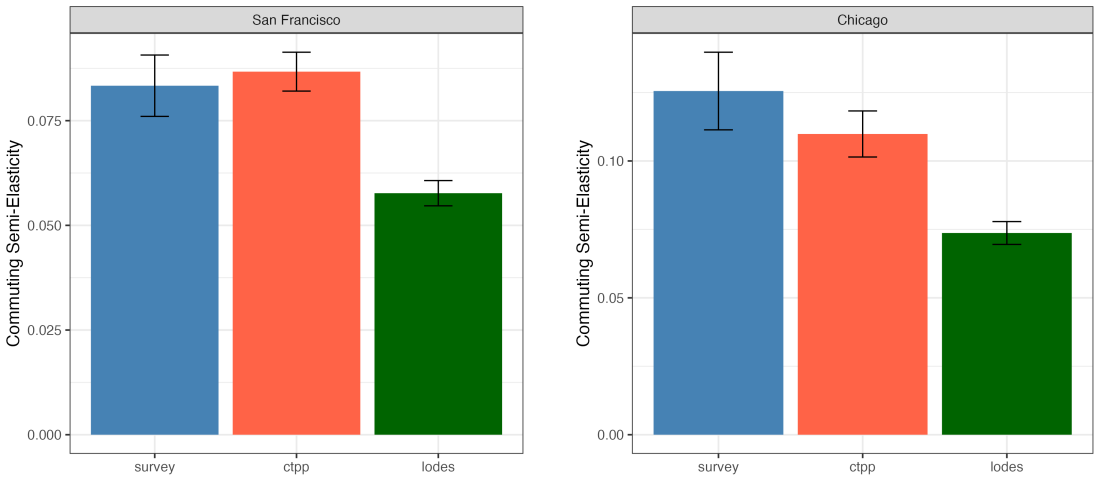
Note: The figures report a version of commuting gravity (2) using PPML estimator for Boston and Portland using CTPP and LODES, where we replace the linear road distance in the dependent variable with dummies of distance bins of 5-10, 10-15, 15-20, 20-25, 25-30, and over 30 kilometers, where we exclude 0-5 kilometer as an omitted baseline category. The shade areas report 95 percent confidence intervals with two-way clustering at the origin and destination level.

Figure A.7: Commuting Semi-Elasticities from CTPP vs LODES



Note: The figures compare the estimates commuting gravity (2) using PPML estimator for each city using CTPP and LODES.

Figure A.8: Gravity Regression of Commuting Flows: Travel Survey vs CTPP vs LODES



(a) San Francisco

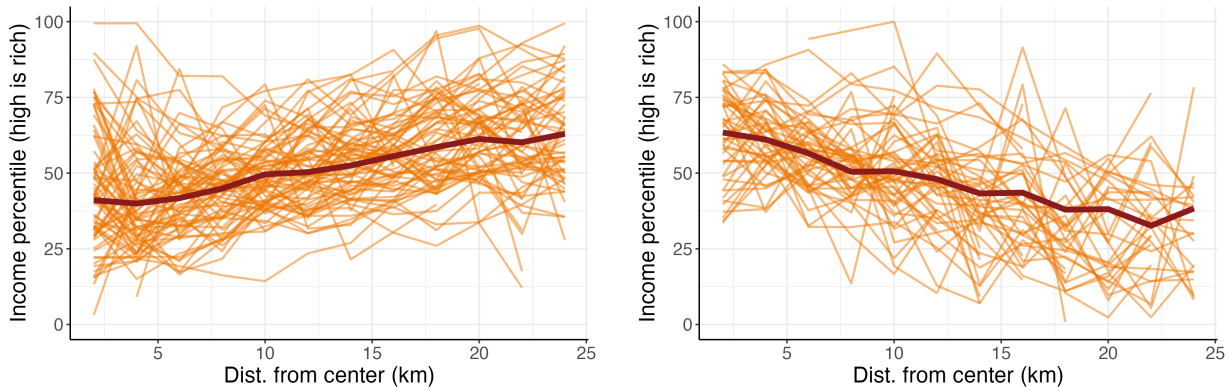
(b) Chicago

Note: The figures compare the estimates commuting gravity (2) using PPML estimator for Chicago and San Francisco, using CTPP (2012-2016), LODES (2015), and travel surveys (2007 for San Francisco and 2000 for Chicago). The error bars report 95 percent confidence intervals with two-way clustering at the origin and destination level.

B. Appendix for Spatial Distribution of Income

B.1. Residential Income and Distance to City Center

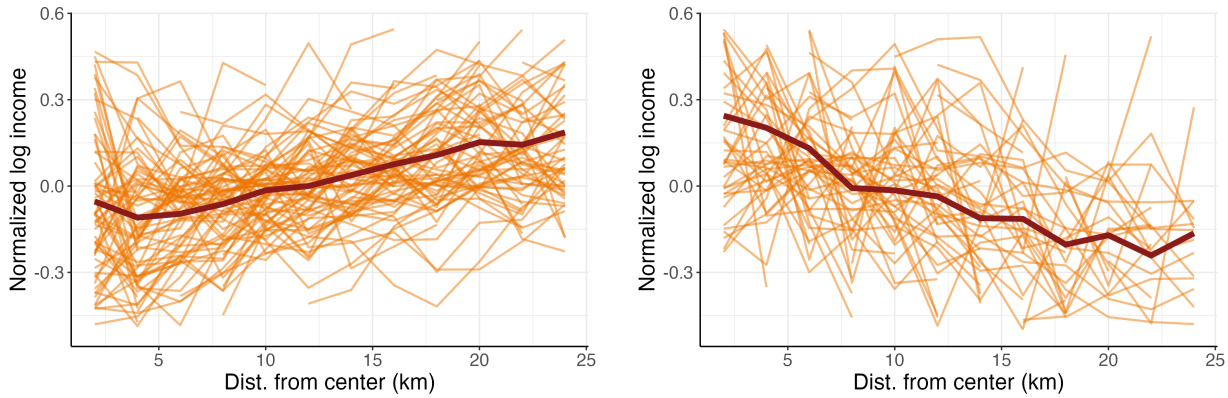
Figure B.1: Residential Income and Distance from City Center



(a) Developed Cities (b) Less Developed Cities

Note: The figures show the relationships between distance from the city center and average neighborhood residential income percentile for developed cities (Panel a) and less-developed cities (Panel b) up to 25 kilometers from the city center. Each light line represents a single city, and averages are highlighted in bold. Appendix Figure 3 shows the same figures for each region.

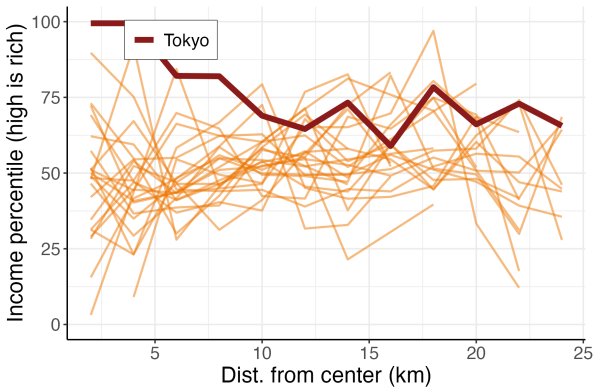
Figure B.2: Residential Income and Distance from City Center: Normalized Log Income



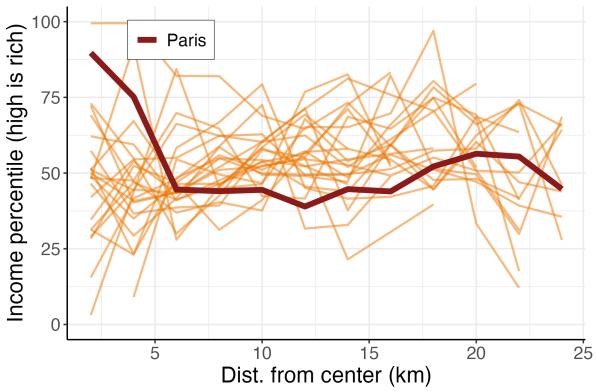
(a) Developed Cities (b) Less Developed Cities

Note: Distance from the city center and income. Emulates Figure 3, except using normalized log income on the y-axis. Each light line represents a single city, and averages are highlighted in bold.

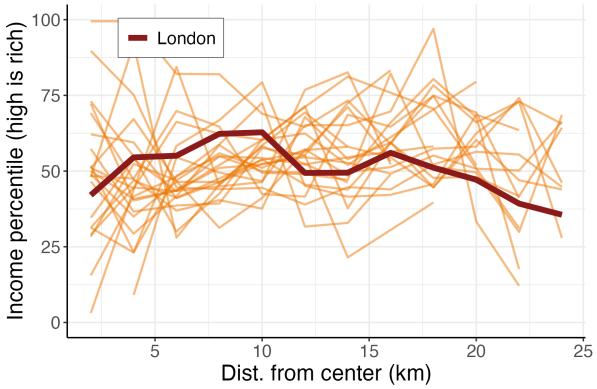
Figure B.3: Residential Income and Distance from City Center: Tokyo, Paris, London



(a) Tokyo



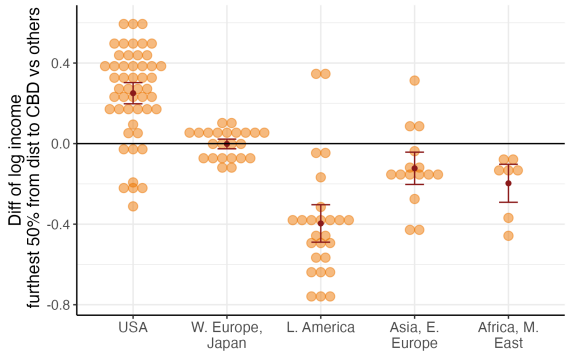
(b) Paris



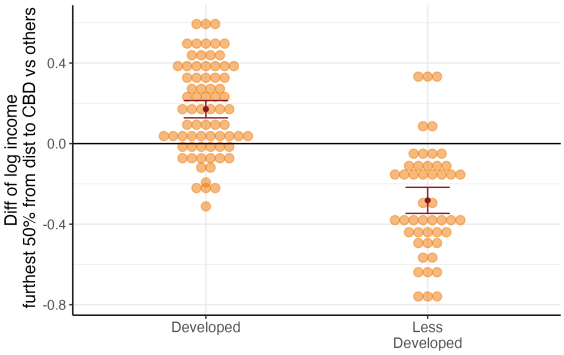
(c) London

Note: The figures show a version of Panel (b) of Figure 3, which shows the income gradient of W. Europe and Japan cities with respect to the city center, highlighting Tokyo, Paris, and London with bold lines.

Figure B.4: Suburban-Urban Income Gap: log Income



(a) By Regions



(b) By Development Status

Note: Emulates Figure 4, except that we take log income, instead of income percentiles within each city.

Table B.1: Suburban-Urban Income Gap: Top and Bottom 20 Cities

City	Country	Difference	Continent	Development Status
Florianópolis	Brazil	-44.4	L. America	Less Developed
Hanoi	Viet Nam	-37.9	Asia, E. Europe	Less Developed
Curitiba	Brazil	-36.0	L. America	Less Developed
Londrina	Brazil	-36.0	L. America	Less Developed
Dhaka	Bangladesh	-35.9	Asia, E. Europe	Less Developed
Porto Alegre	Brazil	-33.9	L. America	Less Developed
Mombasa	Kenya	-33.5	Africa, M. East	Less Developed
Phnom Penh	Cambodia	-33.2	Asia, E. Europe	Less Developed
Tokyo	Japan	-33.2	W. Europe, Japan	Developed
São Paulo	Brazil	-33.0	L. America	Less Developed
Vila Velha	Brazil	-32.4	L. America	Less Developed
Lima	Peru	-31.8	L. America	Less Developed
Belém	Brazil	-28.0	L. America	Less Developed
Uberlândia	Brazil	-26.2	L. America	Less Developed
Fortaleza	Brazil	-25.3	L. America	Less Developed
Rio de Janeiro	Brazil	-24.6	L. America	Less Developed
Seattle	United States	-24.5	USA	Developed
Ho Chi Minh City	Viet Nam	-23.8	Asia, E. Europe	Less Developed
Colombo	Sri Lanka	-21.6	Asia, E. Europe	Less Developed
Cuiabá	Brazil	-21.0	L. America	Less Developed

(a) Bottom 20 Cities

City	Country	Difference	Continent	Development Status
Fresno	United States	39.5	USA	Developed
Bakersfield	United States	37.7	USA	Developed
Las Vegas	United States	37.3	USA	Developed
San Antonio	United States	36.8	USA	Developed
Omaha	United States	35.5	USA	Developed
Indianapolis	United States	33.9	USA	Developed
Buffalo	United States	32.3	USA	Developed
Dà Nang	Viet Nam	31.4	Asia, E. Europe	Less Developed
Oklahoma City	United States	30.8	USA	Developed
Detroit	United States	30.7	USA	Developed
Tucson	United States	30.4	USA	Developed
Columbus	United States	29.6	USA	Developed
Kansas City	United States	28.6	USA	Developed
Albuquerque	United States	28.3	USA	Developed
Providence	United States	27.3	USA	Developed
Cleveland	United States	27.3	USA	Developed
Houston	United States	26.9	USA	Developed
Louisville	United States	25.6	USA	Developed
Sacramento	United States	25.4	USA	Developed
Milwaukee	United States	25.3	USA	Developed

(b) Top 20 Cities

Note: A list of the top and bottom 20 cities, ranked by the gap ratio of income between suburban vs urban areas. Suburban neighborhoods are those containing the furthest 50 percent of the population.

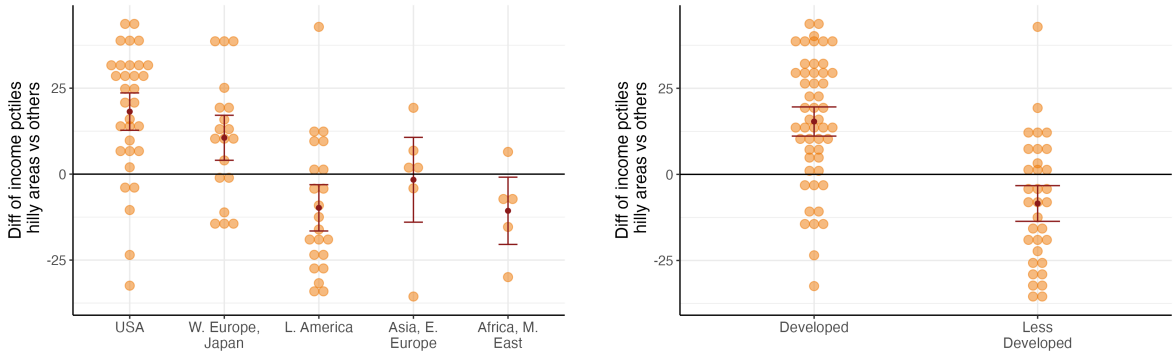
Table B.2: Suburban-Urban Income Gap: By Country

Country	Difference	Cities	continent	Development Status
Bangladesh	-35.9	1	Asia, E. Europe	Less Developed
Cambodia	-33.2	1	Asia, E. Europe	Less Developed
Japan	-33.2	1	W. Europe, Japan	Developed
Peru	-31.8	1	L. America	Less Developed
Sri Lanka	-21.6	1	Asia, E. Europe	Less Developed
D.R. of the Congo	-20.0	1	Africa, M. East	Less Developed
Brazil	-19.9	26	L. America	Less Developed
Kenya	-17.9	2	Africa, M. East	Less Developed
Myanmar	-15.8	1	Asia, E. Europe	Less Developed
Romania	-12.7	1	Asia, E. Europe	Less Developed
Nicaragua	-12.3	1	L. America	Less Developed
Malaysia	-11.7	1	Asia, E. Europe	Less Developed
Syrian Arab Republic	-10.2	1	Africa, M. East	Less Developed
Viet Nam	-10.1	3	Asia, E. Europe	Less Developed
Côte d'Ivoire	-9.9	1	Africa, M. East	Less Developed
Pakistan	-8.6	2	Asia, E. Europe	Less Developed
Lao People's DR	-8.2	1	Asia, E. Europe	Less Developed
Egypt	-7.5	1	Africa, M. East	Less Developed
usa	-7.3	2	USA	Developed
Philippines	-6.3	2	Asia, E. Europe	Less Developed
Indonesia	-5.0	1	Asia, E. Europe	Less Developed
U.R. of Tanzania: Mainland	-4.9	1	Africa, M. East	Less Developed
Spain	0.4	7	W. Europe, Japan	Developed
France	2.5	7	W. Europe, Japan	Developed
United Kingdom	3.9	9	W. Europe, Japan	Developed
China	5.7	1	Asia, E. Europe	Less Developed
United States	18.6	48	USA	Developed

Note: A list of all countries in our neighborhood-level income dataset ranked by the gap ratio of income between suburban vs urban areas. Suburban neighborhoods are those containing the furthest 50 percent of the population.

B.2. Residential Income and Hills/Rivers

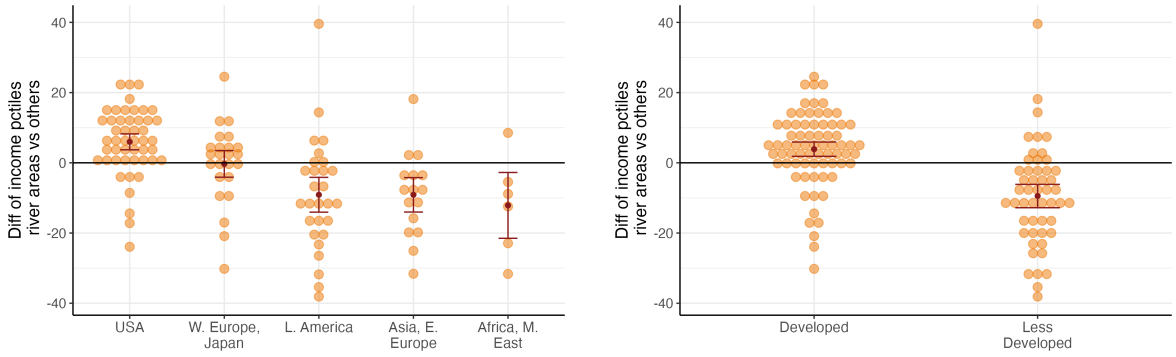
Figure B.5: Residential Income and Hills: By City



(a) By Regions (b) By Development Status

Note: Differences in average income percentile between hilly neighborhoods and non-hilly neighborhoods, as defined in Section 2. Emulates Figure 5, but focuses only on hills. Panel (a) by region; Panel (b) groups cities by development status. We also plot the average value and its 95 percent confidence interval.

Figure B.6: Residential Income and Rivers: By City



(a) By Regions (b) By Development Status

Note: Differences in average income percentile between neighborhoods that are near a river, meaning within 100 meters of a natural waterway. Emulates Figure 5, but focuses only on rivers. Panel (a) by region; Panel (b) groups cities by development status. We also plot the average value and its 95 percent confidence interval.

Table B.3: Residential Income and Hills/Rivers: Top and Bottom 20 Cities

City	Country	Difference	Continent	Development Status
João Pessoa	Brazil	-35.2	L. America	Less Developed
Damascus	Syrian Arab Republic	-32.3	Africa, M. East	Less Developed
Dhaka	Bangladesh	-31.6	Asia, E. Europe	Less Developed
Toulouse	France	-30.2	W. Europe, Japan	Developed
Lima	Peru	-29.2	L. America	Less Developed
Novo Hamburgo	Brazil	-26.0	L. America	Less Developed
Managua	Nicaragua	-25.2	L. America	Less Developed
Lahore	Pakistan	-25.1	Asia, E. Europe	Less Developed
Belém	Brazil	-23.3	L. America	Less Developed
Dar es Salaam	U.R. of Tanzania: Mainland	-22.9	Africa, M. East	Less Developed
Natal	Brazil	-21.8	L. America	Less Developed
Vila Velha	Brazil	-21.5	L. America	Less Developed
Maceió	Brazil	-20.4	L. America	Less Developed
Londrina	Brazil	-20.4	L. America	Less Developed
Colombo	Sri Lanka	-20.3	Asia, E. Europe	Less Developed
Phnom Penh	Cambodia	-19.4	Asia, E. Europe	Less Developed
Newcastle upon Tyne	United Kingdom	-19.3	W. Europe, Japan	Developed
Curitiba	Brazil	-17.7	L. America	Less Developed
Recife	Brazil	-17.4	L. America	Less Developed
Florianópolis	Brazil	-16.8	L. America	Less Developed

(a) Bottom 20 Cities

City	Country	Difference	Continent	Development Status
Campinas	Brazil	39.6	L. America	Less Developed
Honolulu	United States	30.7	USA	Developed
Las Vegas	United States	29.4	USA	Developed
Concord	United States	25.1	USA	Developed
Bradenton	United States	22.9	USA	Developed
Cleveland	United States	22.2	USA	Developed
Dà Nang	Viet Nam	21.9	Asia, E. Europe	Less Developed
Atlanta	United States	21.9	USA	Developed
Los Angeles	United States	21.3	USA	Developed
Austin	United States	21.1	USA	Developed
Marseille	France	19.9	W. Europe, Japan	Developed
San Jose	United States	16.4	USA	Developed
Houston	United States	16.2	USA	Developed
Orlando	United States	15.8	USA	Developed
Málaga	Spain	15.7	W. Europe, Japan	Developed
Buffalo	United States	15.5	USA	Developed
Bakersfield	United States	15.4	USA	Developed
Seattle	United States	15.1	USA	Developed
Sacramento	United States	15.0	USA	Developed
Albuquerque	United States	15.0	USA	Developed

(b) Top 20 Cities

Note: A list of the top and bottom 20 cities, ranked by the gap in income percentile between neighborhoods that are hilly or near a river and thost that are not, as defined in Section 2.

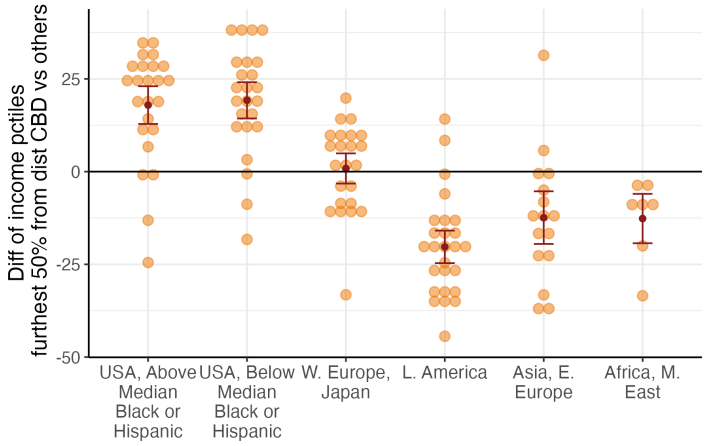
Table B.4: Residential Income and Hills/Rivers: By Country

Country	Ratio	Cities	Continent	Development Status
Syrian Arab Republic	-32.3	1	Africa, M. East	Less Developed
Bangladesh	-31.6	1	Asia, E. Europe	Less Developed
Peru	-29.2	1	L. America	Less Developed
Nicaragua	-25.2	1	L. America	Less Developed
U.R. of Tanzania: Mainland	-22.9	1	Africa, M. East	Less Developed
Sri Lanka	-20.3	1	Asia, E. Europe	Less Developed
Cambodia	-19.4	1	Asia, E. Europe	Less Developed
Pakistan	-14.6	2	Asia, E. Europe	Less Developed
Kenya	-13.9	2	Africa, M. East	Less Developed
Myanmar	-12.3	1	Asia, E. Europe	Less Developed
Japan	-11.3	1	W. Europe, Japan	Developed
Brazil	-10.0	26	L. America	Less Developed
Côte d'Ivoire	-8.9	1	Africa, M. East	Less Developed
Malaysia	-7.5	1	Asia, E. Europe	Less Developed
Lao People's DR	-6.5	1	Asia, E. Europe	Less Developed
Philippines	-5.9	2	Asia, E. Europe	Less Developed
Egypt	-4.4	1	Africa, M. East	Less Developed
Indonesia	-2.8	1	Asia, E. Europe	Less Developed
D.R. of the Congo	-2.1	1	Africa, M. East	Less Developed
France	-1.2	7	W. Europe, Japan	Developed
Viet Nam	-0.3	3	Asia, E. Europe	Less Developed
United Kingdom	1.0	9	W. Europe, Japan	Developed
Romania	1.6	1	Asia, E. Europe	Less Developed
Spain	2.7	7	W. Europe, Japan	Developed
China	2.8	1	Asia, E. Europe	Less Developed
United States	9.5	48	USA	Developed
usa	12.7	2	USA	Developed

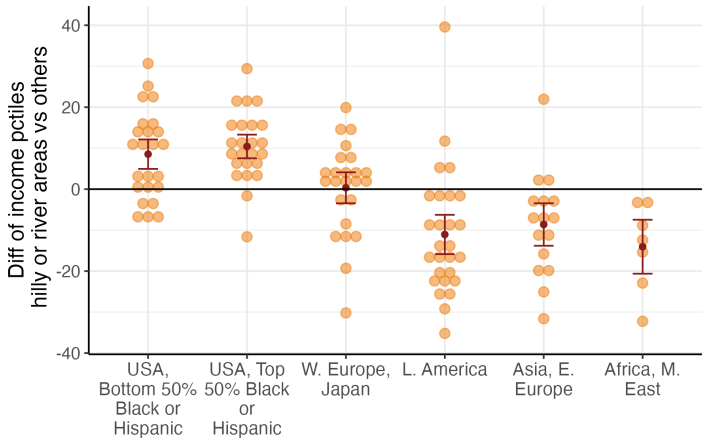
Note: A list of countries, ranked by the gap in income percentile between neighborhoods that are hilly or near a river and those that are not, as defined in Section 2.

B.3. Separating U.S. by Racial Minority Shares

Figure B.7: Residential Income: Separating U.S. by Racial Minority Shares



(a) Suburban-Urban Income Gap



(b) Income Premiums in Hills and Rivers

Note: Emulates Figures 4 and 5, except that we divide the U.S. between cities with above and below median shares of racial minorities (Black or Hispanic).

B.4. Other Statistics

Table B.5: Average Household Size in Less Developed Cities

	Household Size			
	(1)	(2)	(3)	(4)
Suburban	0.06			0.06
	(0.05)			(0.05)
Hilly		0.07		0.06
		(0.05)		(0.05)
River			0.02	0.01
			(0.02)	(0.02)
City Fixed Effects	X	X	X	X
Unique Cities	26	26	26	26
Observations	5,344	5,328	5,344	5,328

*Note: Regression results of average household size in less developed cities. The unit of observation is a neighborhood, and the weights are the inverse number of neighborhoods in each city c. Standard errors are clustered at the city level. ***, ** and * indicate statistical significance at the 1-percent, 5-percent, and 10-percent levels.*

Table B.6: Employment Status in Less Developed Cities

	Working			
	(1)	(2)	(3)	(4)
Suburban	0.01			0.01
	(0.01)			(0.01)
Hilly		−0.01		−0.02
		(0.01)		(0.01)
River			0.01	0.01
			(0.01)	(0.01)
City Fixed Effects	X	X	X	X
Unique Cities	26	26	26	26
Observations	5,344	5,328	5,344	5,328

Note: Regression results of the dummy for working (as opposed to unemployed, non-employed, or students) in less developed cities, analogous to Table B.5.

Table B.7: Average Age in Less Developed Cities

	Average Age			
	(1)	(2)	(3)	(4)
Suburban	−1.29***			−1.21***
	(0.41)			(0.38)
Hilly		−0.19		0.19
		(0.88)		(0.79)
River			−0.83***	−0.55***
			(0.21)	(0.17)
City Fixed Effects	X	X	X	X
Unique Cities	26	26	26	26
Observations	5,344	5,328	5,344	5,328

Note: Regression results of the average age in less developed cities, analogous to Table B.5.

Table B.8: Years of Schooling in Less Developed Cities

	Years of Schooling			
	(1)	(2)	(3)	(4)
Suburban	−0.61**			−0.50**
	(0.19)			(0.16)
Hilly		−0.57		−0.38
		(0.44)		(0.41)
River			−0.47**	−0.33*
			(0.15)	(0.16)
City Fixed Effects	X	X	X	X
Unique Cities	10	10	10	10
Observations	2,273	2,273	2,273	2,273

Note: Regression results of the years of schooling, analogous to Table B.5.

Table B.9: Spatial Patterns of Population Density

Dependent Variable:	Log population density			
Model:	(1)	(2)	(3)	(4)
Variables				
$\text{Developed}_c \times \text{Suburban}_{j,c}$	-0.75***			-0.73***
	(0.06)			(0.06)
$\text{Less Developed}_c \times \text{Suburban}_{j,c}$	-0.93***			-0.82***
	(0.07)			(0.07)
$\text{Developed}_c \times \text{Hilly}_{j,c}$		-0.86***		-0.76***
		(0.13)		(0.10)
$\text{Less Developed}_c \times \text{Hilly}_{j,c}$		-0.55***		-0.48***
		(0.11)		(0.08)
$\text{Developed}_c \times \text{River}_{j,c}$			-0.59***	-0.54***
			(0.06)	(0.05)
$\text{Less Developed}_c \times \text{River}_{j,c}$			-0.74***	-0.57***
			(0.07)	(0.06)
Difference: Less Developed$_c$ vs Developed$_c$				
$\text{Suburban}_{j,c}$	-0.18*			-0.09
	(0.09)			(0.09)
$\text{Hilly}_{j,c}$		0.31*		0.29**
		(0.16)		(0.13)
$\text{River}_{j,c}$			-0.15	-0.03
			(0.09)	(0.08)
Observations	51,551	51,535	51,551	51,535
Unique Cities	125	125	125	125
City FE	✓	✓	✓	✓
Weight by neighborhood pop within city	✓	✓	✓	✓

*Note: Analysis similar to Table 2 using population density, as opposed to income as the outcome variable. Panel (a) reports the results of the regression (1), and the bottom panel reports the differences in the regression coefficients between developed and less-developed cities calculated through at Wald test. Unit of observation is a neighborhood and weights are the inverse number of neighborhoods in each city c. ***, ** and * indicate statistical significance at the 1-percent, 5-percent, and 10-percent levels.*

C. Appendix for Commuting Frictions and Job Access

C.1. Commuting Frictions

Table C.1: Estimated Commuting Semi-Elasticity to Road Distance (km) in Less-Developed Cities

City	Country	Commuting Semi-Elasticity (km)
Chengdu	China	0.69 (0.03)
Phnom Penh	Cambodia	0.66 (0.04)
Lahore	Pakistan	0.60 (0.03)
Hanoi	Viet Nam	0.56 (0.02)
Kinshasa	D.R. of the Congo	0.54 (0.04)
Vientiane	Lao People's DR	0.46 (0.03)
Yangon	Myanmar	0.43 (0.02)
Cebu City	Philippines	0.41 (0.03)
Mombasa	Kenya	0.40 (0.03)
Cairo	Egypt	0.40 (0.02)
Dhaka	Bangladesh	0.39 (0.03)
Manila	Philippines	0.38 (0.03)
Ho Chi Minh City	Viet Nam	0.36 (0.01)
Abidjan	Côte d'Ivoire	0.31 (0.01)
Managua	Nicaragua	0.30 (0.02)
Nairobi	Kenya	0.30 (0.03)
Dar es Salaam	U.R. of Tanzania: Mainland	0.28 (0.02)
Lima	Peru	0.28 (0.01)
Dà Nang	Viet Nam	0.27 (0.04)
Jakarta	Indonesia	0.26 (0.01)
Bucharest	Romania	0.25 (0.01)
Karachi	Pakistan	0.20 (0.02)
Belém	Brazil	0.20 (0.01)
Colombo	Sri Lanka	0.19 (0.01)
Kuala Lumpur	Malaysia	0.17 (0.01)
Damascus	Syrian Arab Republic	0.06 (0.02)

Note: Estimated semi-elasticity of road distance in kilometers using PPML estimator of specification (2)*. Parentheses indicate the standard errors, where the standard errors are clustered in two ways by origins and by destinations.*

Table C.2: Estimated Commuting Semi-Elasticity to Road Distance (km) in W. Europe and Japan

City	Country	Commuting Semi-Elasticity (km)
Marseille	France	0.30 (0.03)
Lyon	France	0.22 (0.01)
Nottingham	United Kingdom	0.21 (0.01)
Leeds	United Kingdom	0.20 (0.01)
Newcastle upon Tyne	United Kingdom	0.20 (0.01)
Liverpool	United Kingdom	0.20 (0.01)
Sheffield	United Kingdom	0.20 (0.01)
Birmingham	United Kingdom	0.20 (0.01)
Toulouse	France	0.19 (0.01)
Manchester	United Kingdom	0.18 (0.00)
Bordeaux	France	0.17 (0.01)
Lille	France	0.17 (0.01)
London	United Kingdom	0.17 (0.00)
Paris	France	0.16 (0.00)
Nice	France	0.15 (0.01)
Bristol	United Kingdom	0.15 (0.02)
Tokyo	Japan	0.14 (0.00)

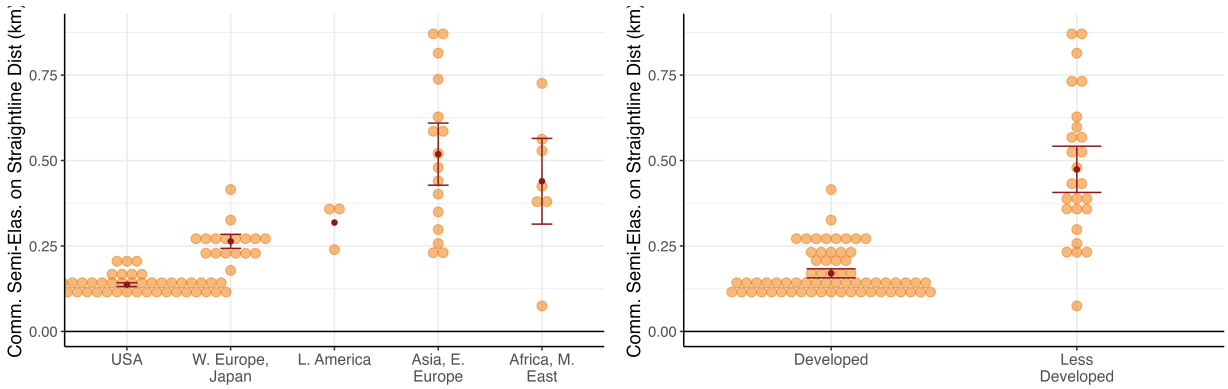
Note: Estimated semi-elasticity of road distance in kilometers using PPML estimator of specification (2). *Parentheses indicate the standard errors, where the standard errors are clustered in two ways by origins and by destinations.*

Table C.3: Estimated Commuting Semi-Elasticity to Road Distance (km) in U.S.

City	Country	Commuting Semi-Elasticity (km)
Pittsburgh	United States	0.17 (0.01)
Boston	United States	0.16 (0.00)
Providence	United States	0.15 (0.01)
Bridgeport	United States	0.14 (0.00)
Philadelphia	United States	0.13 (0.00)
New Orleans	United States	0.13 (0.01)
Portland	United States	0.13 (0.00)
New York	United States	0.12 (0.00)
Tucson	United States	0.12 (0.01)
Virginia Beach	United States	0.12 (0.01)
Buffalo	United States	0.12 (0.01)
Honolulu	United States	0.12 (0.01)
Concord	United States	0.11 (0.01)
Washington D.C.	United States	0.11 (0.00)
St. Louis	United States	0.11 (0.00)
Baltimore	United States	0.11 (0.00)
Minneapolis	United States	0.11 (0.00)
Austin	United States	0.11 (0.00)
Tampa	United States	0.11 (0.00)
Denver	United States	0.11 (0.00)
Atlanta	United States	0.11 (0.00)
Fresno	United States	0.11 (0.01)
Bradenton	United States	0.11 (0.01)
Kansas City	United States	0.11 (0.00)
Cincinnati	United States	0.10 (0.01)
Orlando	United States	0.10 (0.00)
Milwaukee	United States	0.10 (0.00)
Louisville	United States	0.10 (0.01)
Indianapolis	United States	0.10 (0.00)
Sacramento	United States	0.10 (0.00)
Chicago	United States	0.10 (0.00)
San Antonio	United States	0.10 (0.00)
Albuquerque	United States	0.10 (0.01)
Omaha	United States	0.10 (0.01)
Cleveland	United States	0.10 (0.00)
Oklahoma City	United States	0.10 (0.00)
Dallas	United States	0.10 (0.00)
Seattle	United States	0.10 (0.00)
Salt Lake City	United States	0.09 (0.01)
Columbus	United States	0.09 (0.00)
Los Angeles	United States	0.09 (0.00)
Bakersfield	United States	0.09 (0.01)
Las Vegas	United States	0.09 (0.00)
Miami	United States	0.09 (0.00)
San Jose	United States	0.09 (0.00)
Detroit	United States	0.08 (0.00)
Houston	United States	0.08 (0.00)
Phoenix	United States	0.08 (0.00)

Note: Estimated semi-elasticity of road distance in kilometers using PPML estimator of specification (2). Parentheses indicate the standard errors, where the standard errors are clustered in two ways by origins and by destinations.

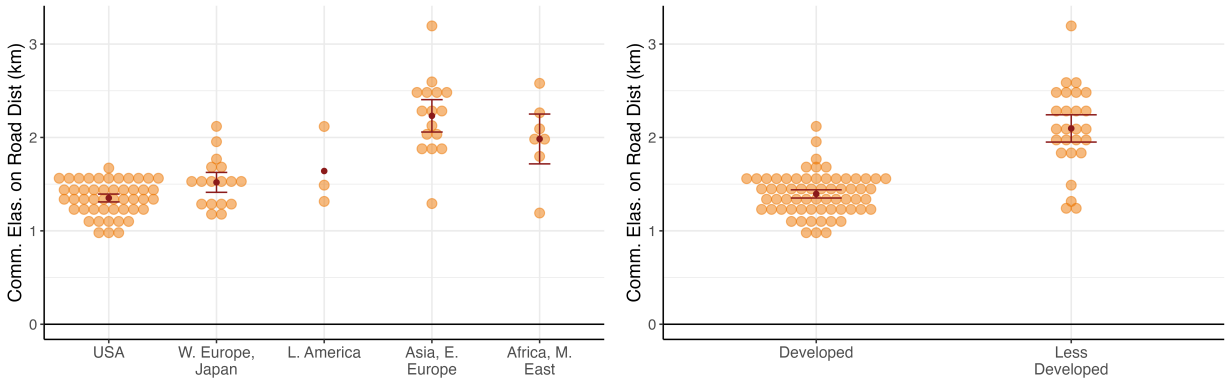
Figure C.1: Estimated Semi-Elasticity of Commuting to Straightline Distance



(a) By Regions (b) By Development Status

Note: A version of Figure 6, where we replace road distance with straightline distance.

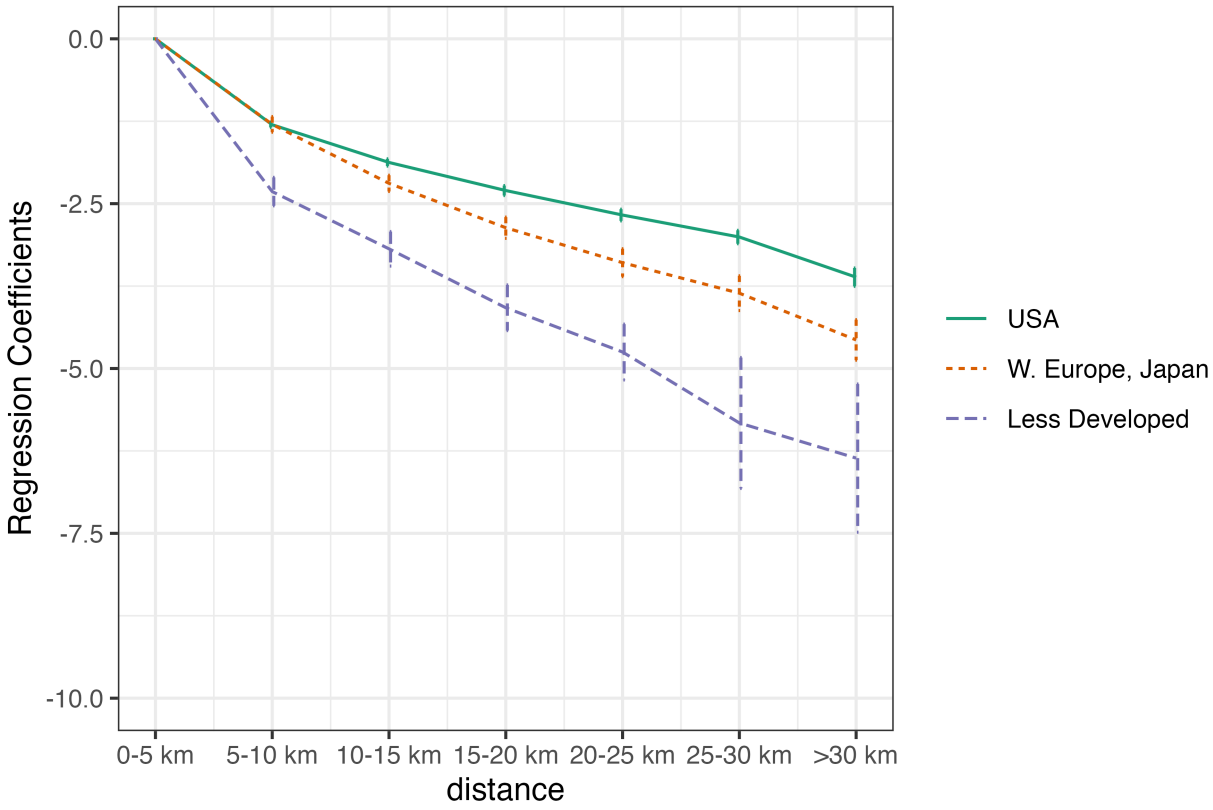
Figure C.2: Estimated Elasticity of Commuting to Road Distance



(a) By Regions (b) By Development Status

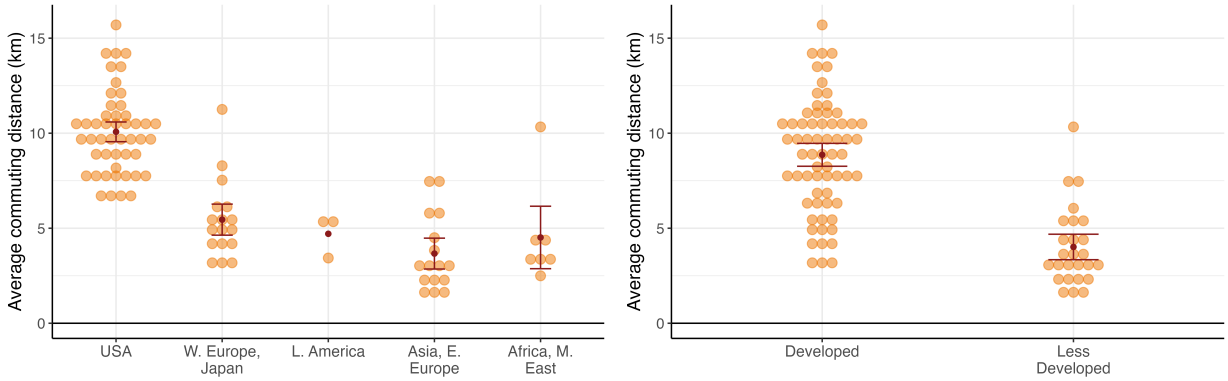
Note: A version of Figure 6, where we replace the road distance in the dependent variable with its log transformation.

Figure C.3: Nonparametric Gravity Regression Results



Note: Estimated results of a version of specification (2) using PPML estimator, where we replace the linear road distance in the dependent variable with dummies of distance bins of 5-10, 10-15, 15-20, 20-25, 25-30, and over 30 kilometers, where we exclude 0-5 kilometer as an omitted baseline category. We estimate this specification for each city, and present the mean and 90 percent confidence intervals of the mean estimates within each of the three sets of countries.

Figure C.4: Average Commuting Distances



(a) By Regions

(b) By Development Status

Note: The figures show the average road distances between the home and work locations. Each dot represents an estimate for each city. Panel (a) groups them by region, while Panel (b) groups cities by development status.

Table C.4: Commuting Semi-Elasticity, Speed, and Modes

	Dependent variable:		
	log Commuting Semi-Elasticity		
	(1)	(2)	(3)
Speed Index (log scale)	-1.83^{***}		-0.30
	(0.14)		(0.33)
Prvivate Car		-1.29^{***}	-1.11^{***}
		(0.08)	(0.22)
Constant	-1.75^{***}	-1.00^{***}	-1.11^{***}
	(0.04)	(0.06)	(0.13)
Observations	71	71	71
R^2	0.72	0.80	0.80

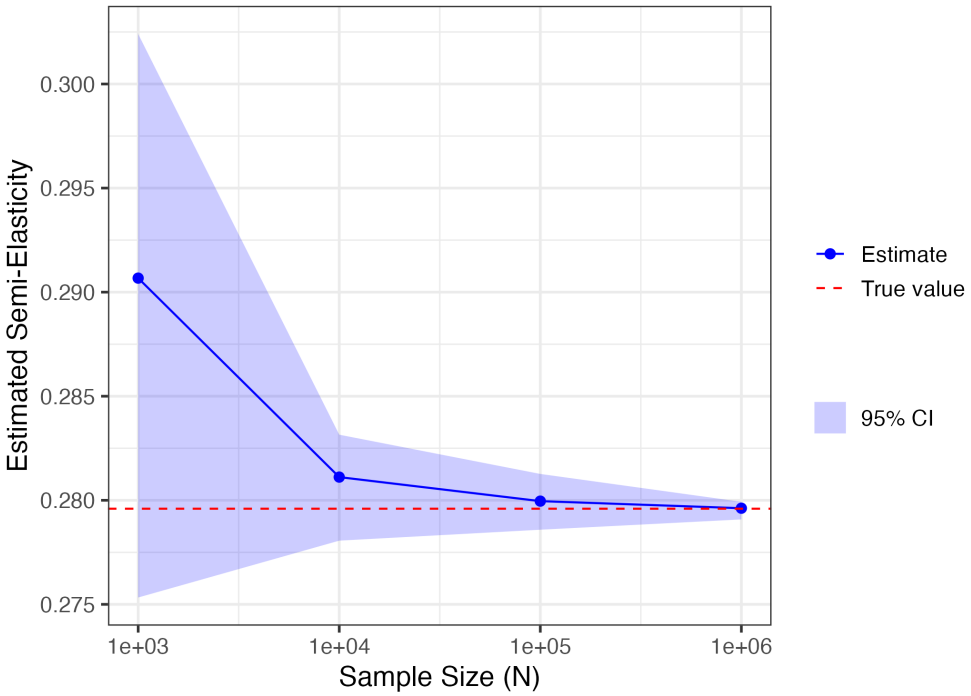
This table shows the regression results of the estimated semi-commuting elasticities from specification (2) on "speed index" from Akbar et al. (2023b), which measures log-point differences in average road speed across cities using Google Maps API, and on the fraction of commuters by private car. We measure the latter value using our travel surveys whenever they are available, and for other cities, we use the data constructed by Prieto-Curiel and Ospina (2024) based on various surveys and administrative data. We restrict to cities where both of these data sets are available.

C.2. Small Sample Properties of Commuting Semi-elasticity Estimates

In this appendix, we assess how the sampling rates of travel surveys may affect the estimates of commuting semi-elasticities through a Monte Carlo simulation. Specifically, we assume that the estimated commuting gravity (2) from Lima (with 388 survey zones) is the "ground truth, and we simulate random samples following this data generation process. We vary the number of commuters to draw ($N = 10^3, 10^4, 10^5, 10^6$). For each N, we sum up all commuters across all location pairs, and estimate the same commuting gravity equation (2). We repeat this process 20 times for each N and draw a 95% confidence interval of the estimated semi-elasticities.

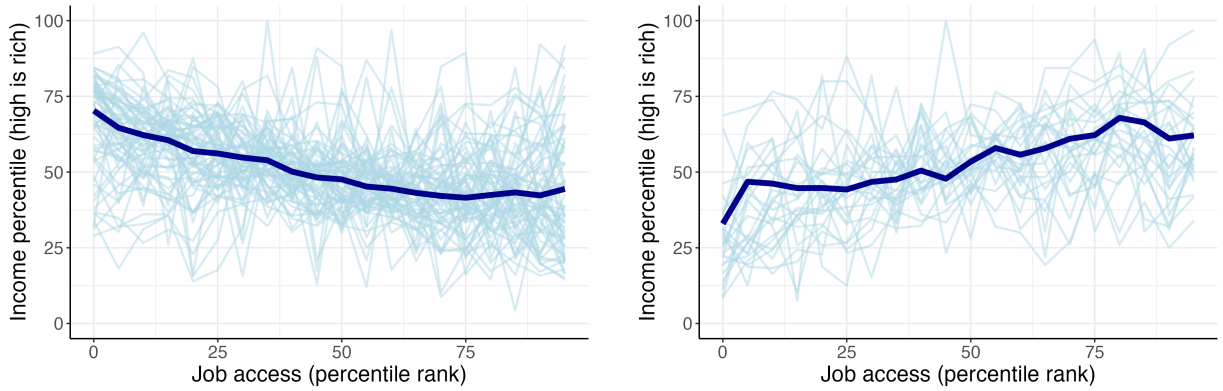
We find that the statistical uncertainty quickly diminishes as the sample size increases. Even with $N = 1000$, the mean is 0.29 relative to the assumed true value of 0.28, with confidence intervals ranging rightly from 0.27 to 0.30. With $N = 10000$, the bias and variance become almost negligible. Note that the actual sample size of household travel surveys from Lima is approximately 66,000 (excluding students), which is roughly the same as the average sample size of other travel surveys. Therefore, we conclude that the statistical uncertainty associated with sample size is negligible.

Figure C.5: Commuting Semi-Elasticities from CTPP vs LODES



C.3. Job Access

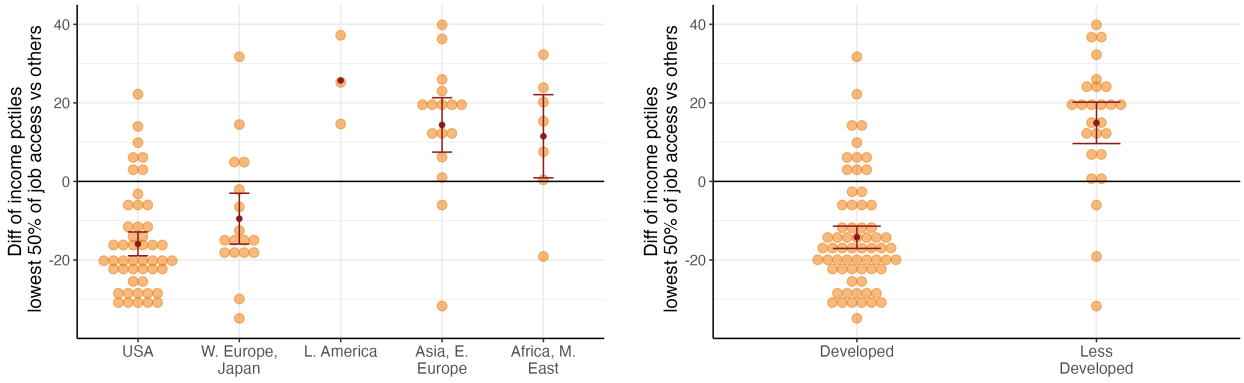
Figure C.6: Residential Income and Job Access, by Development Status



(a) Developed Cities (b) Less Developed Cities

The figures show the relationships between the percentiles of estimated job access (equation 3) within each city and average neighborhood residential income percentile for developed cities (Panel a) and less-developed cities (Panel b), instead of by region in Figure 8. Each light line represents a single city, and averages are highlighted in bold.

Figure C.7: Residential Income and High-Job-Access Neighborhoods



(a) By Regions (b) By Development Status

Note: The figures display the difference in average income percentiles between high and low job access neighborhoods for each city, where high job access is defined by above median job access neighborhoods. Each dot represents a city. Panel (a) groups them by region, while Panel (b) groups them by development status. In both panels, we also report the group averages and their 95 percent confidence intervals.

Table C.5: Regression Results of Job Access and Work Location Fixed Effects

Dependent Variables:	Job access (home FE)	Area-adjusted work location FE
Model:	(1)	(2)
Variables		
$\text{Developed}_c \times \text{Suburban}_{j,c}$	-0.64***	-0.69***
	(0.04)	(0.06)
$\text{Less Developed}_c \times \text{Suburban}_{j,c}$	-1.6***	-1.8***
	(0.14)	(0.13)
$\text{Developed}_c \times \text{Hilly}_{j,c}$	-0.19***	-0.94***
	(0.05)	(0.13)
$\text{Less Developed}_c \times \text{Hilly}_{j,c}$	-0.82***	-1.1***
	(0.23)	(0.22)
$\text{Developed}_c \times \text{River}_{j,c}$	-0.09***	-0.15***
	(0.02)	(0.04)
$\text{Less Developed}_c \times \text{River}_{j,c}$	-0.50***	-0.67***
	(0.11)	(0.13)
Difference: Less Developed$_c$ vs Developed$_c$		
$\text{Suburban}_{j,c}$	-0.95***	-1.1***
	(0.14)	(0.14)
$\text{Hilly}_{j,c}$	-0.62***	-0.19
	(0.23)	(0.25)
$\text{River}_{j,c}$	-0.40***	-0.52***
	(0.11)	(0.13)
Observations	36,029	35,937
Unique Cities	91	91
City FE	✓	✓

Note: The results of regression (1), where the outcome variables are estimated origin fixed effects from the commuting gravity equation (2) (Column 1); and destination fixed effect, net of log area (Column 2), where the area adjustment. For Column 2, the adjustment is applied to offset the mechanical effect that the destination fixed effects proportionally increase in neighborhood area size (Kreindler and Miyauchi, 2023). In our quantitative model, Column (1) corresponds to expected wage rates at residential location $\overline{w}_{j,c}$, and Column (2) corresponds to productivity $A_{j,c}$.

D. Model Fit

Table D.1 demonstrates that the estimated model closely replicates key patterns in the spatial distribution of income across U.S. cities. The table presents a version of regression (1), where the odd-numbered columns report results using model-predicted log residential income, and the even-numbered columns use observed income data. In the independent variable, $\overline{w}_{j,c}$ is the estimated job access in Section 4. We instrument $\overline{w}_{j,c}$, $\overline{w}_{j,c} \times \text{Hilly}_{j,c}$ and $\overline{w}_{j,c} \times \text{River}_{j,c}$ by $\tilde{\overline{w}}_{j,c}$, $\tilde{\overline{w}}_{j,c} \times \text{Hilly}_{j,c}$, and $\tilde{\overline{w}}_{j,c} \times \text{River}_{j,c}$, where $\tilde{\overline{w}}_{j,c}$ is constructed analogously as in Section 4, except we replace actual road distances for the estimation of equation (2) with bilateral straight-line distances, consistent with the moment condition used for the model estimation (Section 6.1).

Columns (1) and (2) show that the model successfully replicates the elevated average residential income in suburban, hilly, and river-adjacent neighborhoods. Columns (3) and (4) capture the unconditional negative relationship between residential income and job access, consistent with the pattern illustrated in Section 4. In Columns (5) and (6), once we condition on suburban, hilly, and river dummies, both the model and the data exhibit a positive association between residential income and job access, both in the model and data. Finally, Columns (7) and (8) show that the interaction between job access and geographic features is negative in both the model and the data. While the model's regression coefficients do not exactly match those from the data – due in part to overidentifying restrictions in the GMM estimation – the overall patterns are well aligned.

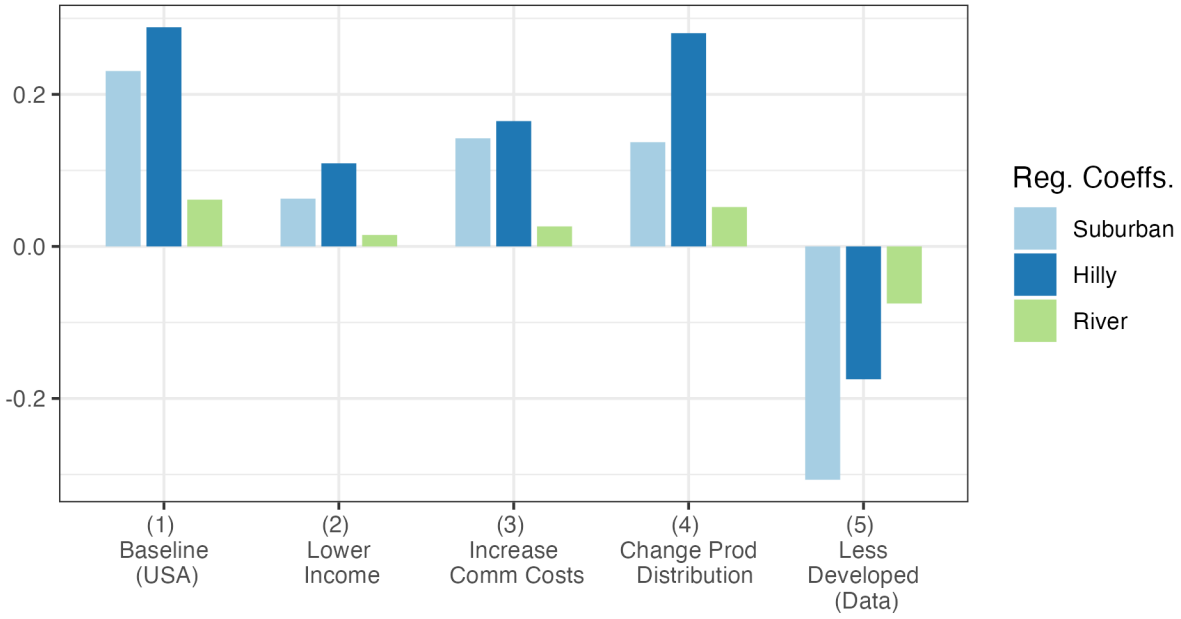
Table D.1: Model Fit to U.S. cities

	log Residential Income$_{j,c}$							
	Model	Data	Model	Data	Model	Data	Model	Data
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Suburban$_{j,c}$	0.23	0.23			0.29	0.27	0.29	0.27
	(0.01)	(0.03)			(0.01)	(0.04)	(0.01)	(0.04)
Hilly$_{j,c}$	0.29	0.29			0.31	0.30	0.32	0.31
	(0.01)	(0.05)			(0.01)	(0.06)	(0.01)	(0.06)
River$_{j,c}$	0.06	0.07			0.08	0.07	0.08	0.08
	(0.004)	(0.02)			(0.003)	(0.02)	(0.003)	(0.02)
$\overline{w}_{j,c}$			−0.41	−0.56	0.55	0.35	0.45	0.27
			(0.08)	(0.17)	(0.05)	(0.20)	(0.05)	(0.20)
$\overline{w}_{j,c} \times$ Hilly$_{j,c}$							0.51	0.33
							(0.08)	(0.30)
$\overline{w}_{j,c} \times$ River$_{j,c}$							0.22	0.17
							(0.03)	(0.15)
City Fixed Effects	X	X	X	X	X	X	X	X
Unique Cities	48	48	48	48	48	48	48	48
Observations	27,117	27,117	27,117	27,117	27,117	27,117	27,117	27,117

Note: A version of regression (1), with model-predicted log residential income in odd columns, and actual data in even columns. In the independent variable, $\overline{w}_{j,c}$ is as estimated in Section 4, using the value of $\theta = 5$. We instrument $\overline{w}_{j,c}$, $\overline{w}_{j,c} \times Hilly_{j,c}$ and $\overline{w}_{j,c} \times River_{j,c}$ by $\tilde{\overline{w}}_{j,c}$, $\tilde{\overline{w}}_{j,c} \times Hilly_{j,c}$, and $\tilde{\overline{w}}_{j,c} \times River_{j,c}$, where $\tilde{\overline{w}}_{j,c}$ is constructed analogously as in Section 4, except we replace actual road distances for the estimation of equation (2) with bilateral straight-line distances. Standard errors are clustered at the city level.

E. Additional Figures and Tables for Quantitative Analysis

Figure E.1: Spatial Residential Income Distribution of U.S. Cities: Separate Counterfactual for Lowering Income, Increasing Commuting Costs, and Changing Productivity Distribution



Note: This figure displays the estimated coefficients on the suburban, hilly, and river indicators from regression specification (1). Column (1) presents the baseline results from our model calibrated to U.S. cities, corresponding to Column (1) of Table D.1. Column (2) shows the regression coefficients under a counterfactual in which average productivity levels $\{A_{j,c}\}_j$ are uniformly reduced by 2.0 log points across all neighborhoods and cities—approximately 14% of their baseline values. Column (3) reports coefficients from a counterfactual equilibrium in which we alternatively increase the commuting semi-elasticity κ_c by 0.14, reflecting the average gap between U.S. and less-developed cities as shown in Figure 6. Column (4) presents results from an alternative counterfactual in which the relative productivity of suburban, hilly, and river neighborhoods is adjusted to match the patterns observed in less-developed cities (Table C.5). Finally, Column (5) displays the corresponding estimates for less-developed cities based on observed data, as reported in Table 2.

Table E.1: Estimated $\{\beta_1, \beta_2, \beta_3\}$ to Rationalize the Income Distribution in Other Developed Cities

Model	Suburban (β_1)	Hilly (β_2)	River (β_3)
Baseline Estimates (USA)	0.82	1.14	0.26
Estimates to Fully Rationalize Non-U.S. Developed Cities	0.51	0.70	0.19

Note: First row reports the estimated $\{\beta_1, \beta_2, \beta_3\}$ using U.S. cities as reported in Table 4. Second row reports the estimated $\{\beta_1, \beta_2, \beta_3\}$ to fully rationalize the differences in the income premiums of suburban, hilly, and river neighborhoods, after accounting for the overall productivity differences, commuting cost differences, and differences in productivity premiums in those areas, as described further in Section 6.2.

Table E.2: Spatial Variations of Equivalent Variations (EV) from Uniform Productivity Increase

	Dependent variable:			
	EV (High Earning Potential)		EV (Low Earning Potential)	
	(1)	(2)	(3)	(4)
Suburban		-0.056^{***}		-0.075^{***}
		(0.003)		(0.004)
Hilly		-0.053^{***}		-0.050^{**}
		(0.016)		(0.024)
River		-0.010^{***}		-0.017^{***}
		(0.002)		(0.002)
High Work Access		0.023^{***}		0.025^{***}
		(0.003)		(0.003)
Constant	1.884^{***}	1.921^{***}	1.883^{***}	1.897^{***}
	(0.003)	(0.004)	(0.002)	(0.004)
Number of cities	48	48	48	48
Observations	27,117	27,117	27,117	27,117
R^2	0.000	0.511	−0.000	0.568

Note: Regression of EVs, in log points, across households by residential location and earning potential, to increase the overall productivity from low-income-city level to high-income-city level (as observed in the U.S. cities). "High Work Access" proxies the neighborhoods with above-median work access for each city.

F. Heterogeneity of Commuting Costs and Spatial Productivity Distribution by Income Groups

In our baseline analysis, we abstracted from income-related heterogeneity in commuting costs and wage distributions. However, in U.S. cities, disparities in transportation modes and commuting costs across income groups have been identified as one potential driver of residential sorting by income (Glaeser et al., 2008; Su, 2022). In this subsection, we assess the quantitative relevance of this channel.

To do so, we extend our model from Section 5 to allow commuting costs and wages to depend on individuals' earning potential s. Specifically, we now let commuting costs $\tau_{jn,c}(s)$ and wages $w_{n,c}(s)$ vary with s. We retain the assumption that idiosyncratic preferences over work locations are drawn from an i.i.d. Fréchet distribution with shape parameter θ, which is common across s. Under this setting, the probability that a worker with earning potential s living in neighborhood j commutes to job location n is given by:

$$\lambda_{jn,c}(s) = \frac{\left(\tau_{jn,c}(s)^{-1} w_{n,c}(s)\right)^{\theta}}{\sum_{\ell}\left(\tau_{j\ell,c}(s)^{-1} w_{\ell,c}(s)\right)^{\theta}}. \tag{F.1}$$

We also follow Section 4 that the commuting cost $\tau_{jn,c}$ is a power function of road distance between neighborhoods j to n such that $\log \tau_{jn,c}(s) = \kappa_c(s) \times \text{RoadDistance}_{jn,c}$, where $\kappa_c(s)$ can depend on s. Under this assumption, the commuting gravity equation (11) holds separately for each earning potential s:

$$\log \lambda_{jn,c}(s) = \theta \log w_{n,c}(s) - \kappa_c(s)\theta \times \text{RoadDistance}_{jn,c} - \log \sum_{\ell}\left(\tau_{j\ell,c}(s)^{-1} w_{\ell,c}(s)\right)^{\theta}. \tag{F.2}$$

If we observe the commuting flows by earning potential s, we can estimate the empirical analog of this equation using a PPML estimator for each earning potential s, analogously as equation (2):

$$\log \mathbb{E}\left[\lambda_{jn,c}(s)\right] = \psi_{n,c} - \kappa_c(s)\text{RoadDistance}_{jn,c} + \eta_{j,c}(s), \tag{F.3}$$

where $\psi_{n,c}(s) \equiv \theta \log w_{n,c}(s)$ is the workplace fixed effects, $\kappa_c(s) \equiv \kappa_c(s)\theta$, and $\eta_{j,c}(s) \equiv -\log \sum_{\ell}\left(\tau_{j\ell,c}(s)^{-1} w_{\ell,c}(s)\right)^{\theta}$ is the origin fixed effects.

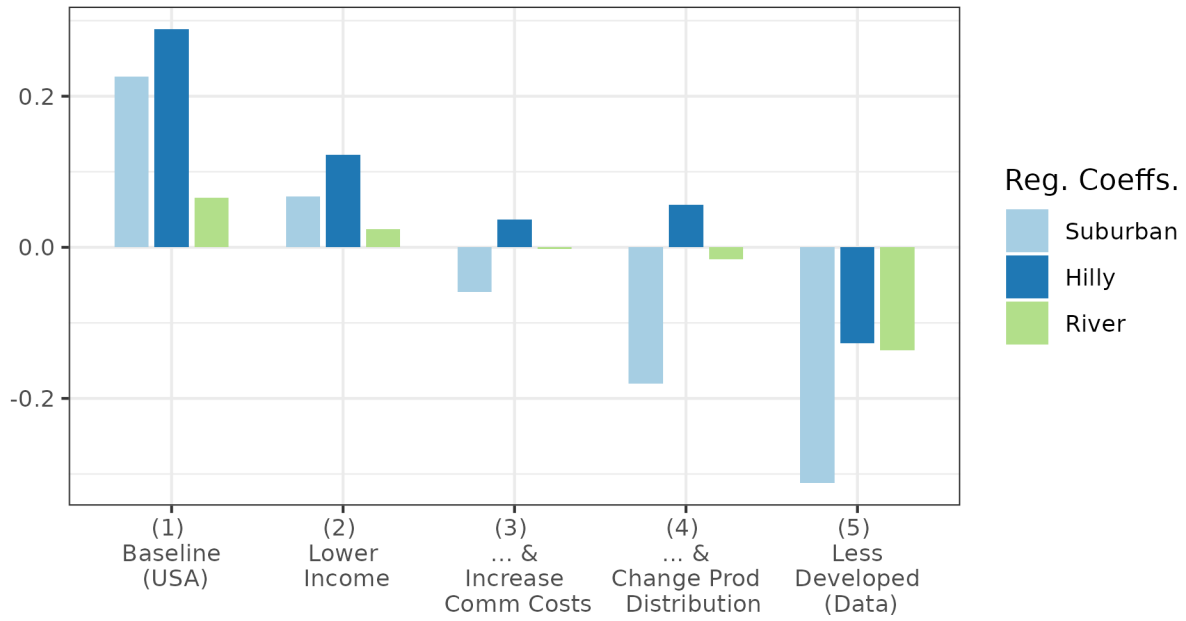
A challenge in estimating this equation is that we do not directly observe earning potentials s in our data. While some travel survey data provide information about variables such as education, it is not comprehensive, and it is also difficult to create a proxy that is consistent across countries. To deal with this issue, we divide our samples based on the realized income. Through the lens of our model, realized income is a product of earning potential s and the expected wage rates at the residential location $\overline{w}_{j,c}$. Therefore, if the variance of earning potential s is much larger than the spatial variation of wage rates $\overline{w}_{j,c}$, this strategy effectively splits the samples with high- and low-earning-potential households within each city.

Figure 11 reports the estimated commuting semi-elasticities for developed cities (U.S. cities and Tokyo; commuting data by income are not available for the U.K. and France) and for less-

developed cities. Consistent with previous studies, we find larger commuting semi-elasticities among low-income groups (0.11) than high-income groups (0.10) in developed cities. However, this difference is *more* pronounced in less-developed cities, with 0.42 for low-income groups than 0.32 for high-income groups. These findings suggest that, while these forces may contribute to the sorting of low-income households toward central areas in both developed and less-developed cities, they are unlikely to account for the observed disparity between the two groups of cities.

Using this calibrated model, we run the same counterfactual simulations as before, except that we change the commuting costs and spatial productivity distribution for each s. Consistent with the observations above, we find that changing commuting costs and productivity distributions to the levels of less-developed cities generates slightly smaller effects on the spatial income distribution compared to our baseline scenario (Figure 10). Therefore, the heterogeneity across earning potential s in commuting costs and spatial productivity distribution does not explain the differences in spatial income distribution. At the same time, our main conclusion remains: differences in aggregate income levels, commuting costs, and spatial productivity distributions *across cities* jointly account for the main observed gap in spatial income distributions between U.S. and less-developed cities.

Figure F.1: Unpacking Spatial Income Distribution Gap: U.S. vs Less-developed Cities, Using Model with Heterogeneous Commuting Costs and Wage Distribution across Earning Potential



Note: A version of Figure 10, where we use the extended model to incorporate heterogeneity in commuting costs and spatial productivity distribution across s. See Table F.1 for the differences in the estimated neighborhood productivity by income levels $\{A_{j,c}(s)\}_j$.

Table F.1: Regression Results of Estimated A_j and $\overline{w}_j$ by Income Groups

Dependent Variables:	Area-adj Work FE (High Earner)	Area-adj Work FE (Low Earner)	Job Access (High Earner)	Job Access (Low Earner)
Model:	(1)	(2)	(3)	(4)
Variables				
$\text{Developed}_c \times \text{Suburban}_{j,c}$	-0.60***	-0.60***	-0.60***	-0.59***
	(0.06)	(0.07)	(0.05)	(0.05)
$\text{Less Developed}_c \times \text{Suburban}_{j,c}$	-1.5***	-1.6***	-1.3***	-1.4***
	(0.12)	(0.11)	(0.09)	(0.10)
$\text{Developed}_c \times \text{Hilly}_{j,c}$	-1.0***	-1.0***	-0.19***	-0.19***
	(0.16)	(0.15)	(0.06)	(0.06)
$\text{Less Developed}_c \times \text{Hilly}_{j,c}$	-0.87***	-1.0***	-0.73***	-1.0***
	(0.32)	(0.34)	(0.23)	(0.31)
$\text{Developed}_c \times \text{River}_{j,c}$	-0.30***	-0.36***	-0.15***	-0.16***
	(0.05)	(0.05)	(0.02)	(0.02)
$\text{Less Developed}_c \times \text{River}_{j,c}$	-0.61***	-0.64***	-0.41***	-0.48***
	(0.10)	(0.10)	(0.09)	(0.10)
Difference: Less Developed$_c$ vs Developed$_c$				
$\text{Suburban}_{j,c}$	-0.93***	-0.99***	-0.70***	-0.82***
	(0.14)	(0.13)	(0.10)	(0.11)
$\text{Hilly}_{j,c}$	0.16	-0.005	-0.54**	-0.84***
	(0.36)	(0.38)	(0.24)	(0.31)
$\text{River}_{j,c}$	-0.32***	-0.28***	-0.26***	-0.32***
	(0.11)	(0.11)	(0.09)	(0.10)
Observations	31,229	31,609	31,677	31,677
Unique Cities	74	74	74	74
City FE	✓	✓	✓	✓
Weight by neighborhood pop within city	✓	✓	✓	✓
Subset	✓	✓	✓	✓

Note: The results of regression (1), where the outcome variables are estimated destination fixed effect from the commuting gravity equation (F.3), net of log area and scaled by $\theta = 5$ (Columns 1-2) and the estimated origin fixed effects (Columns 3-4). Odd columns report the results of high-income groups, and even columns report the results of low-income groups, where high- and low-income groups are defined by above or below the city-specific median income.