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

Peter Deffebach, David Lagakos, Yuhei Miyauchi, Eiji Yamada

June 14, 2025

Abstract

How does the spatial distribution of income in cities vary with income per capita? We address this question using new granular evidence from 127 cities in 26 developed and less-developed countries. We document that household income levels decline significantly more rapidly with distance to city centers in less-developed countries. Urban neighborhoods with natural amenities—particularly hills and rivers—are poorer than average in less-developed nations but richer than average in developed ones. We evaluate potential explanations for these patterns using a quantitative spatial model disciplined by commuting gravity equations that we estimate for each city. We find that the spatial income patterns are largely explained by nonhomothetic preferences over amenities, coupled with higher commuting costs and spatially more concentrated jobs in less developed cities. Our model predicts unequal welfare gains across households as citywide income increases.

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, 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 the NBER Summer Institute. 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, Ryan Duswalt, Maria Mittelbach, Isaac Otoo, 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, 2023, and the references therein). The literature still lacks a more comprehensive picture on 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 covering 145,000 neighborhoods in 127 cities of all levels of income per capita. These data, which we describe in detail below, come from a mix of commuting surveys and fine-grained census data that allow us to document how residential locations, commuting patterns, and job access vary by household income within cities at different levels of development. The data cover a number of the world's largest Asian and African cities, as well as middle-income mega cities like Lima and Sao Paolo, and prominent examples from the developed world, like Los Angeles, London, Paris, and Tokyo. Our broad goal is to use these data to understand how residential and commuting decisions of richer and poorer households in cities differ across the world income distribution.

We begin by documenting several prominent ways in which the income distribution within cities varies between developed and developing countries. First, we show that in less developed cities, average residential income decreases steadily with distance to city centers. In developed-world cities, in contrast, income-distance gradients are generally flat or increasing. 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 jobs on average than richer ones.

The second fact we document relates to how average incomes within urban areas vary with the presence of natural amenities. We focus on proximity to hills and waterways, which are natural features of the landscape that are typically thought to improve the residential value of a neighborhood. Our data show that in less developed cities, average incomes are significantly lower than average in hilly neighborhoods or those in close proximity to waterways. The opposite is true in developed cities, where average incomes are substantially *higher* in both hilly areas and those near water. Thus, to the extent that richer households value residential amenities more than poorer ones, less developed cities appear to get desired location outcomes exactly backward, with poorer people placed systematically further from jobs but closer to amenities that they likely value less than their richer counterparts.

What explains these starkly different patterns between cities in less-developed and developed countries? We conduct a series of robustness exercises to provide evidence against a number of appealing explanations related to sample, measurement, and potentially omitted variables. We find that our patterns hold, for example, even when excluding the United States, in which a majority of cities have positive income-distance gradients, or Brazil, with steeply downward sloping income-distance gradients. The patterns also hold when restricting attention only to cities of similar sizes in terms of population or square kilometers, or when capping the outer boundaries of all cities at cutoffs like 15km or 20km. Our facts also cut across the old-world, new-world divide (Henderson, Squires, Storeygard, and Weil, 2018), and are not driven by an "east-side story" (Heblich, Trew, and Zylberberg, 2021). Having ruled out these alternative explanations, we turn to theoretical mechanisms that could explain the patterns we observe.

Our analysis is based on a quantitative spatial model, which has become the gold standard for understanding spatial patterns of economic activity at a fine level of disaggregation (Redding and Rossi-Hansberg, 2017; Redding, 2023). We incorporate five basic potential explanations into the model. The first is non-homothetic preference over amenities, such that richer households care more about neighborhood amenities like those available in leafy suburbs far from the city center, or the views from hilly areas or of water. The second is the comparatively worse transportation infrastructure, and slower commuting speeds, of less developed countries (Akbar, Couture, Duranton, and Storeygard, 2023a,b; Tsivanidis, 2023). The third is more centralized jobs in developing countries (Baum-Snow, 2020; Davis and Dingel, 2020), which we allow as a possibility in the model by having a steeper productivity decline in distance from city center. The fourth is heterogeneity in commuting costs across income groups (LeRoy and Sonstelie, 1983; Glaeser, Kahn, and Rappaport, 2008; Su, 2022), with greater heterogeneity in less developed countries, where car ownership is more concentrated among richer households. Finally, the model allows for the possibility that amenity gradients decline faster with distance from city center, or are relatively worse in hills and rivers, in less developed cities. This can interact with non-homothetic preferences in ways that cause higher-income households to prefer more central locations in developed countries and avoid natural amenities in less developed ones.

In our model, ex-ante heterogeneous households with different earning potential decide residential locations based on job access, housing costs, and neighborhood amenities. We posit non-homothetic preferences over housing and amenities, using a tractable and parsimonious formulation in which neighborhood amenities can be potentially more or less desired relative to consumption and housing as household income rises. Our preferences imply that in equilibrium, the high housing costs of "attractive" neighborhoods push out poorer residents. What constitutes an "attractive" neighborhood fundamentally depends on the overall income levels of residents. In cities with low average incomes, the primary consideration of households is access to jobs. As income rises, residents place a higher value on neighborhood amenities. Therefore, which neighborhoods are more attractive changes as the city's overall income level rises.

To quantify the relevance of these channels in explaining the gap in spatial income distributions between developed and less developed cities, we begin by measuring the commuting costs city by city through estimation of our model's commuting gravity equation, which predicts that commuting flows depend on bilateral commuting costs and destinations' wages. Our data allow for these calculations in a large number of U.S. cities, using the LODES data, Tokyo, and 26 cities in the developing world. These surveys come from the Japan International Cooperation Agency (JICA), which commissioned them as part of urban transportation projects in partner countries. The surveys report residential and workplace locations for a comprehensive sample of on average 70,000 urban residents, along with demographic, employment, and income characteristics.

Our estimates point to significant heterogeneity in commuting costs across cities at different development levels. While the semi-elasticity of commuting with road distances (km) is only 0.07 in US cities and 0.11 in Tokyo, those in less developed cities are substantially higher, ranging from 0.10 to 0.30. These differences lead to disproportionately lower commuting access in suburban areas and those with natural amenities in less developed cities compared to developed cities. Importantly, when estimating these commuting costs by income groups within each city, we find that the heterogeneity in commuting costs is an order of magnitude larger across cities than within cities across income groups, suggesting a modest quantitative explanatory power of commuting cost heterogeneity by household income level within a city.

The commuting gravity equations also help summarize the rich information in our data on how concentrated job opportunities are in city centers, relative to more suburban locations, and how that varies with development. Similarly, they are also informative about the relative prevalence of jobs in areas with natural amenities. In particular, the destination fixed effects from our gravity regressions capture the extent to which each neighborhood in each city is a commuting destination, and hence an area with job opportunities. When comparing suburban areas to city centers, we show that less developed cities have job opportunities that are modestly more concentrated in city centers relative to cities in developed nations. In nearly all cities in our database, areas with natural amenities have relatively low concentrations of job opportunities.

We use the gravity regressions and various other moments of the data to estimate the full equilibrium of our model and quantify how much each potential explanation can account for the observed patterns of spatial income distribution. To do so, we first calibrate our model to match data from U.S. cities. We then simulate the counterfactual effects of changing each of the following to the level of a less developed city: (1) overall city income (productivity), which is relevant because of the model's non-homothetic preferences for amenities; (2) commuting costs, and (3) the distribution of productivity across space. We also simulate the effects of allowing for heterogeneity in commuting cost by income, but relegate this to an extension section, since our empirical analysis suggests it explains very little. We leave differential amenity gradients by development as an unexplained category due to lack of direct data on amenities and their valuations by neighborhood and by city.

We find that when lowering the overall city income, 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 that they also become less attractive residential locations when overall income levels fall. 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 three-quarters of the observed gaps in income premiums in suburban, hilly, and river neighborhoods between the U.S. cities and less-developed cities.

While lower incomes, higher commuting costs, and more spatially concentrated jobs explain much of the observed gap in spatial income distributions between developed and less developed cities, some residual variation remains. These unexplained patterns likely reflect differences in amenity gradients within cities by development status. Central cities 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 (McCulloch, Schaelling, Turner, and Kitagawa, 2025), and rivers in less developed cities are almost surely dirtier than their counterparts in richer countries. We leave the task of measuring and quantifying the importance of the channels to future work.

We conclude by examining how these patterns of spatial income distribution shape the unequal welfare gains associated with overall city development. 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 household types.

Our paper contributes to a growing body of work studying how cities in developed and developing countries systematically differ. 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.[1] 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

[1] See, for example, Tsivanidis (2023), Zárate (2024), Balboni, Bryan, Morten, and Siddiqi (2020), and Khanna, Nyshadham, Ramos-Menchelli, Tamayo, and Tiew (2023) for transportation infrastructure in Bogota, Mexico City, Dar es Salaam, and Medellin, respectively; Michaels, Nigmatulina, Rauch, Regan, Baruah, and Dahlstrand (2021) for residential infrastructure in Dar es Salaam; Franklin, Imbert, Abebe, and Mejia-Mantilla (2024) for urban public works program in Dar es Salaam, respectively; and Harari and Wong (2021) and Gechter and Tsivanidis (2023) for slum upgrading interventions in Jakarta and Mumbai, respectively. Coeurdacier, Oswald, and Teignier (2022) study the structural transformation of land use in French cities since 1870.

and public access and public goods provision within Brazil and India, 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. To our knowledge, no prior paper has documented how the spatial distribution of income and commuting patterns vary across cities over the development spectrum.

We also contribute to the empirical and theoretical literature on the spatial distribution of income within cities that has largely focused on developed countries, and in particular the United States. This work has reconciled income sorting patterns within U.S. cities using non-homotheticities in demand for housing or land (Alonso, 1964; Becker, 1965; Margo, 1992; Hoelzlein, 2023; Finlay and Williams, 2022; Couture, Gaubert, Handbury, and Hurst, 2024), transportation infrastructure (Glaeser et al., 2008; Su, 2022), and natural amenities (Lee and Lin, 2018). Some authors study differences between USA and other developed countries and seek the explanations in nonhomotheticity in amenities (Brueckner, Thisse, and Zenou (1999) and Almagro and Domínguez-Iino (2024) for European cities and Tabuchi (2019) for Tokyo).

2. Data

This section outlines how we integrate diverse data sources to build a comprehensive database on household income, neighborhood characteristics, and commuting flows for 127 cities across 26 countries.

2.1. Travel Surveys from Developing Countries

Our primary data source for developing countries is a collection of microdata from 30 household travel surveys conducted in 26 cities across 21 countries. These surveys, which are central to urban and transportation planning, typically 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 contrast, in developing countries, they are often conducted by city governments on an ad hoc basis – typically in preparation for major infrastructure projects or city master plans – and often with support from international aid agencies.

We compile household-level travel surveys conducted or supervised by the Japan International Cooperation Agency (JICA) as part of efforts to design urban transportation improvements.

The resulting microdata covers 26 cities in 21 developing countries. 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). In four cities (Dhaka, Ho Chi Minh City, Nairobi, and Phnom Penh), we have access to multiple survey waves conducted in different years.

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 approximately 70,000 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 km^2.

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 is 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 aligns with the census data in both the relative ranking of neighborhoods by income 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

Table 1: List of Cities Covered in Our Analysis

	L. America	Asia, E. Europe	Africa, M. East
Number of City-Years	3	19	8
Number of Cities	3	16	7
Number of Countries	3	12	6
Avg Number of Respondents	70469	69774	42873
Avg Number of Neighborhoods	190	206	166
List of Cities	Belem (00), Lima (03), Managua (98)	Bucharest (98), Cebu (14), Chengdu (00), Colombo (13), Da Nang (08), Dhaka (09), Dhaka (14), Hanoi (05), Ho Chi Minh (03), Ho Chi Minh (14), Jakarta (18), Karachi (11), Kuala Lumpur (99), Lahore (10), Manila (96), Phnom Penh (00), Phnom Penh (12), Viang Chan (07), Yangon (13)	Abidjan (13), Cairo (01), Damascus (98), Dar es Salaam (07), Kinshasa (18), Mombasa (15), Nairobi (05), Nairobi (13)

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

	USA	W. Europe, Japan	L. America	Asia, E. Europe	Africa, M. East
Number of Cities	48	24	32	16	7
Number of City-Years	48	24	33	19	8
Number of Countries	1	4	3	12	6
Total Number of Neighborhoods	27655	18027	94447	3930	1334
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 continent

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 survey data, and Figure A.1 for a map.

respondent, the survey typically records whether the individual is employed (either as a wage 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 USA, France, Spain, UK, Brazil We collect average residential neighborhood income from census and tax data for four developed countries (USA, Spain, France, UK) and a less-developed country (Brazil). Information from the United States stems from the 2008-2012 aggregated American Community Survey at the level of census tract. For French cities, income is derived from tax returns at the level of "IRIS", with an average size one quarter that of a census block.[2] In the UK, average income from the Office of Tax Statistics is available at the "Small Area" level, which is slightly larger than a census block. In Spain, average neighborhood income derived from tax information is available at the *Sección*-level. In Brazil, average neighborhood income from the 2010 census is available at the *Setores*-level, with an average size one-tenth that of a US census block.

Travel Survey from Tokyo, Japan For Tokyo, Japan, we have access to the microdata of 2018 Tokyo Person Trip Survey (Tokyo Metropolitan Area Transportation Planning Council, 2018). The data shares a similar structure as 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.

Commuting Flows in USA We obtain commuting flow information in the USA from the Longitudinal Employer-Household Dynamics Origin Destination Employment Statistics (LODES) dataset from the year 2015. The LODES data reports the aggregate number of workers living and working in any given pair of census tracks, as well as living-working flows disaggregated by a coarse measure of income: workers earning \$0-\$30k, \$60-\$90k, and \$90+.

Housing Rents in USA We obtain median housing rents for U.S. cities from the 2008-2012 aggregated American Community Survey (ACS) at the level of census tract. We use this information for calibrating our quantitative model in Section 6.

[2] We use median neighborhood income, instead of mean income, for France, as the latter is not publicly available.

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 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.[3] Although this approach is heuristic, it aligns well with intuitive notions of a city center. 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, an open-source algorithm for finding the shortest path between two locations along OSM's road network (Luxen and Vetter, 2011).

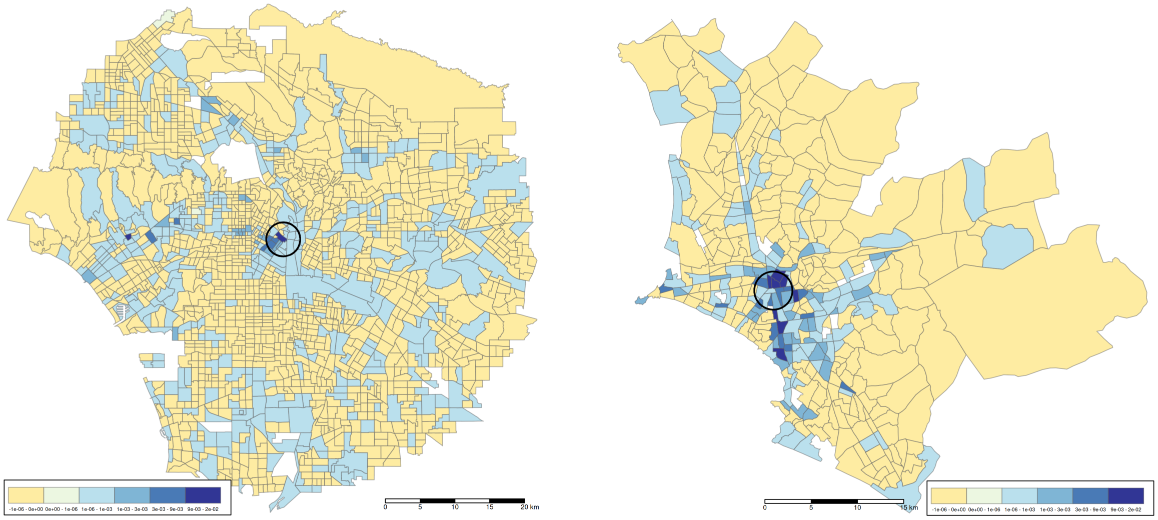
Hills We classify a neighborhood as *hilly* if its average slope exceeds 5 degrees, based on 30m×30m elevation data from Amazon Web Services Terrain Tiles (Larrick, Tian, Rogers, Acosta, and Shen, 2020).[4]

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

[3]For example, Cebu City's center is defined at the city hall, while London's is at Trafalgar Square.

[4]Slope is calculated as the average change in elevation across four adjacent grid cells (Hijmans, 2024). 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: 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.

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.

Development Status We define *developed* cities as those in the USA, Spain, France, UK, and Japan. We define *less-developed* cities as those surveyed by JICA and Brazil. The poorest country in our data is Kinshasa, Democratic Republic of the Congo, with a GDP per capita of 1,020 USD.

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 US, UK, Spain, France, and Brazil are derived from the underlying administrative data source directly.

2.4. Final Data Sets

Our final neighborhood-level income dataset includes 145,000 neighborhoods across 127 cities in 26 countries. 72 are classified as "developed" and 55 are classified as "less-developed". For each neighborhood, we observe average household income, distance to the city center, and its geographic features. Panel (b) of Table 1 lists all the cities in our analysis. Appendix Figure A.1 shows a map of all cities in our data across the world.

Our final pairwise commuting flow dataset includes all 30 JICA cities, all 48 cities in the USA, and Tokyo. For each home-work pair, we observe the proportion of the working population living and working in that pair. We also disaggregate pairwise commuting flows by coarse measures of income.

3. Spatial Distribution of Income

This section documents the distinctive patterns that characterize the spatial distribution of income in developed versus less-developed cities.

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

Before turning to the full statistical analysis, we begin with an illustrative comparison between two cities: Los Angeles, USA, a developed city, and Lima, Peru, a less-developed city.

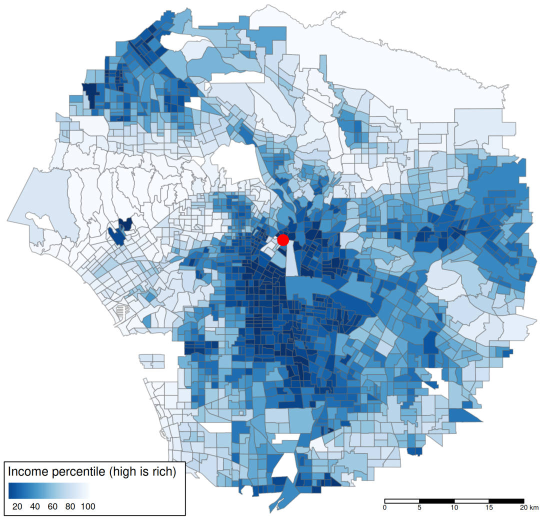
Figure 2 illustrates income distribution and geographic features for Los Angeles and Lima. Panels (a) and (b) show the average residential income by neighborhood, measured as each neighborhood's percentile rank within the city, with lighter colors corresponding to higher income levels. The red dot in 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.

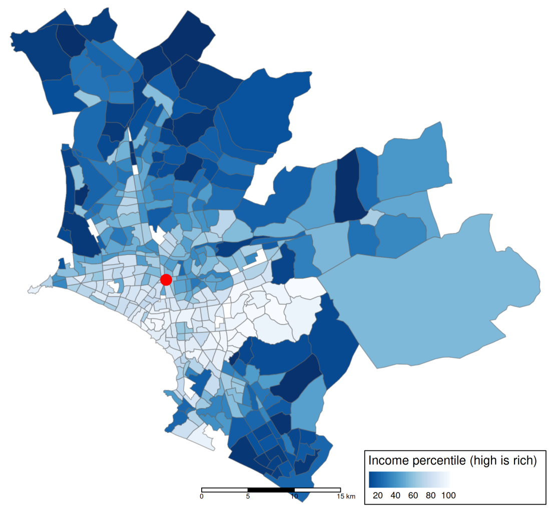
Panels (c) and (d) reveal similarly contrasting patterns between income and hilly areas. In Los Angeles, hilly areas such as Laurel Canyon and Beverly Hills – located northwest of the city center – are associated with high incomes. In Lima, by contrast, Los Olivos, a middle-income area nestled in a valley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

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

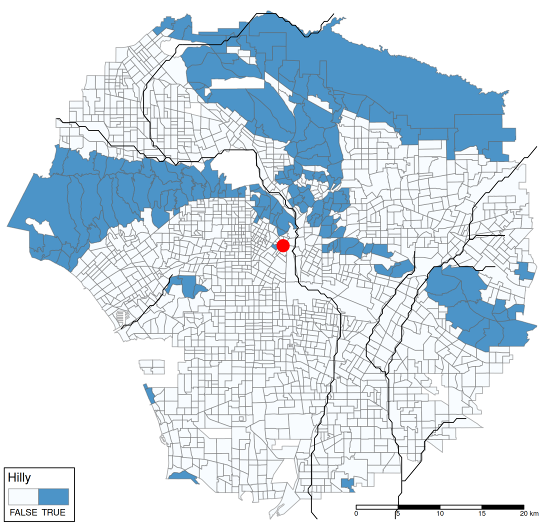
Figure 2: Residential Income and Hilly Areas in Los Angeles and Lima



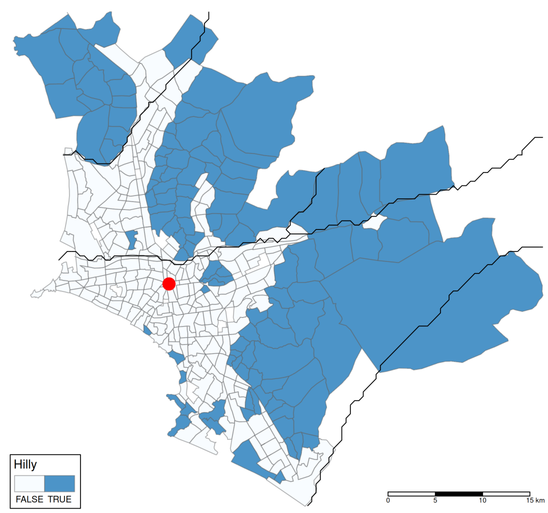
(a) Residential Income (Los Angeles)



(b) Residential Income (Lima)



(c) Hills and Rivers (Los Angeles)



(d) Hills and Rivers (Lima)

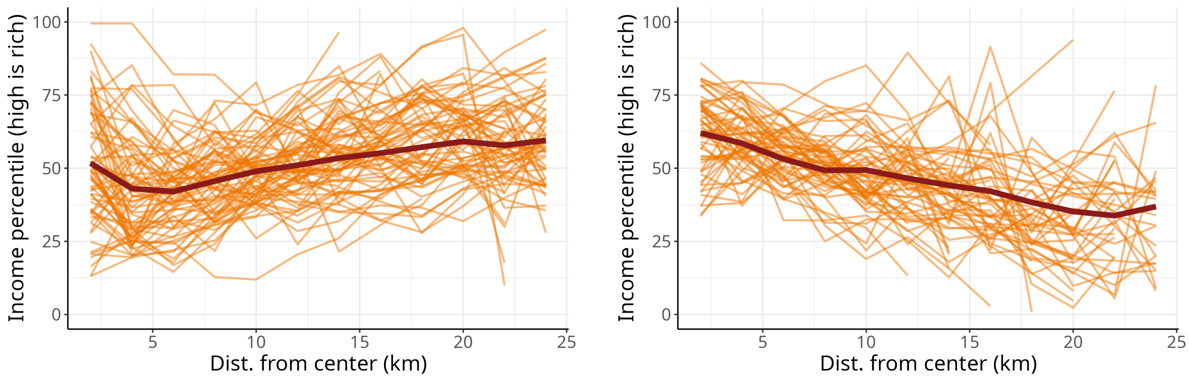
Note: Panels (a) and (b) show the average residential income by neighborhood, measured as each neighborhood's percentile rank within the city, with lighter colors corresponding to higher income levels. Figures (c) and (d) show a binary measure for hilliness (blue is hilly) along with the path of waterways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

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 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, while averages are highlighted in bold.

Figure 3: Residential Income and Distance from City Center



(a) Developed Cities (b) Less Developed Cities

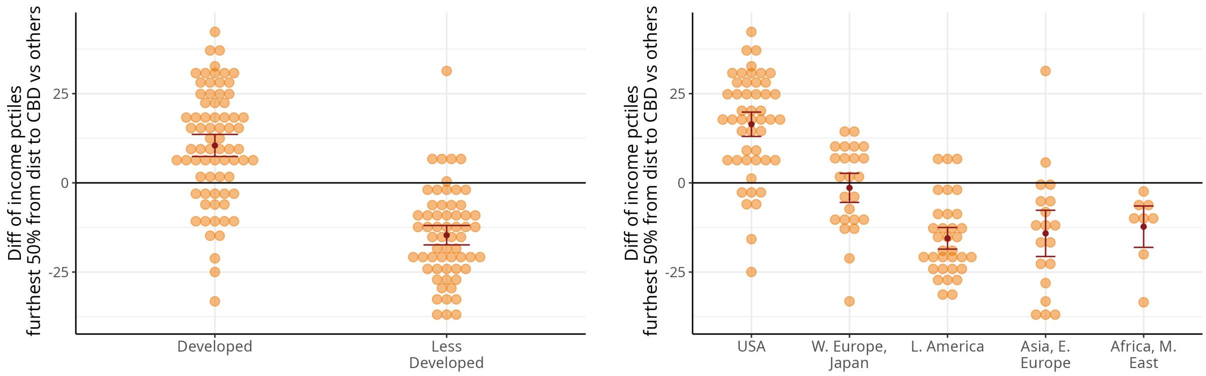
Note: The figures show the relationships between distance from the city center and average neighborhood residential income percentile for developed cities (Panel a) and less-developed cities (Panel b) up to 25 kilometers from the city center. Each light line represents a single city, and averages are highlighted in bold. Appendix Figure B.1 shows the same figures for each continent. Appendix Figure B.2 shows the same figures using normalized log income instead of income percentiles.

Panel (a) shows that in developed 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 relatively affluent, with income levels slightly above the 50th percentile. Income declines sharply just outside the center, reaching a low point at around the 40th percentile at approximately 5 kilometers from the city center. Beyond this point, income gradually increases toward the urban periphery, approaching the 60th percentile. Overall, this pattern reflects a positive income gradient from the inner suburbs to the outer edges of the city.

In contrast, Panel (b) reveals a strong, monotonic decline in income with distance from the city center in less-developed cities. Central neighborhoods average around the 60th percentile in income, but this steadily falls to approximately the 35th percentile at 25 kilometers from the center, indicating a pronounced negative gradient.

Figure 4 further examines cross-city variation in the relationship between income and distance

Figure 4: Suburban-Urban Income Gap



(a) By Development Status

(b) By Continents

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 cities by development status, while Panel (b) groups them by continent. In both panels, we also report the group averages and their 95 percent confidence intervals. Appendix Figure B.3 shows the same set of figures using in log income instead of income percentiles. Appendix Tables B.1 and B.2 report the values for each city and country, respectively.

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.[5] Panel (a) groups cities by development status, while Panel (b) disaggregates by continent. In both panels, each dot represents a city, and we plot group means along with 95 percent confidence intervals.

Panel (a) reveals a stark contrast between developed and less-developed cities. In developed cities, the suburban–urban income gap is positive, averaging around 10 percentile points, indicating that suburban areas tend to be richer than central areas. In contrast, less-developed cities exhibit a negative gap of approximately 15 percentile points, with suburban areas systematically poorer than their urban cores. While the gaps vary across cities, the average difference between the two groups is both large and statistically significant.

Panel (b) highlights variation in these patterns across continents. Among developed regions, the suburban–urban income gap is most pronounced in the United States (around 15 percentile points), while it is close to zero in Western Europe and Japan, suggesting relatively flat income

[5] Results are robust to alternative definitions of suburban areas (e.g., using the outermost 25 percent of the population) and to alternative distance measures; see Appendix Figure B.4 and Table B.5.

gradients with respect to distance from the center (see Appendix Figure B.1 for detailed line plots). Among less-developed regions, the gap is consistently negative across Latin America, Asia, Eastern Europe, Africa, and the Middle East.

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, USA, are from developed countries. Conversely, 19 of the 20 cities with the largest positive gaps are in developed countries – all in the United States – 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 United States, 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.

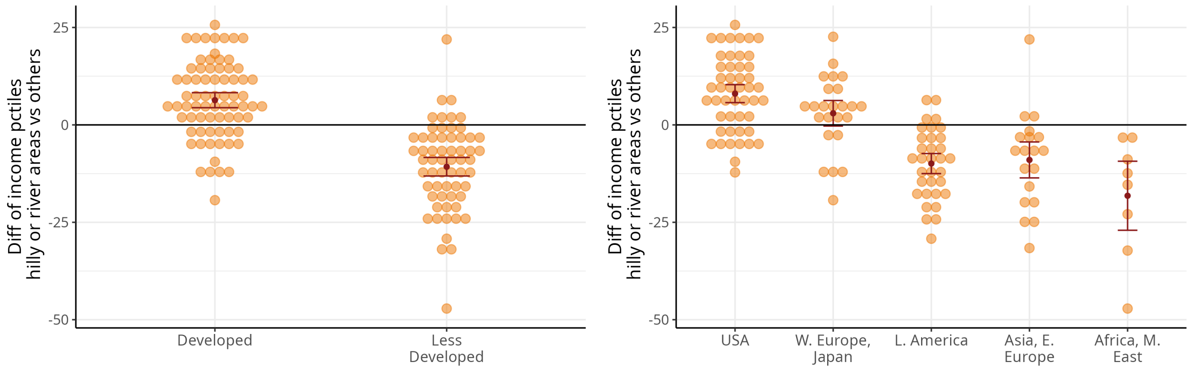
Figure 5 presents the difference in average income percentiles between neighborhoods located in hilly or river-adjacent areas and those that are not, city by city. For simplicity, we combine hilly and river neighborhoods in this figure. In the regression analysis that follows, we show that the patterns remain qualitatively similar when hills and rivers are analyzed separately.

Panel (a) shows that, on average, neighborhoods located in hilly or river-adjacent areas are approximately 10 percentile points richer than other neighborhoods in developed cities, and about 5 percentile points poorer in less-developed cities. Both differences are statistically significant. Panel (b) highlights heterogeneity within developed regions: the income gap is largest in the United States, while in Western Europe and Japan, the average gap is smaller, though still positive and marginally statistically significant. Unlike the suburban–urban income gap documented in Figure 4, even Western Europe and Japan exhibit a modest income premium in hilly or river-adjacent areas.

Among less-developed cities, the pattern is reversed. Hilly or river neighborhoods tend to be significantly poorer, with negative income gaps across Latin America, Asia and Eastern Europe, and Africa and the Middle East.

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, 19

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



(a) By Development Status

(b) By Continents

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 cities by development status; Panel (b) groups them by continent. 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.

are located in less-developed countries, with only one from a developed country (Newcastle, United Kingdom). Conversely, among the 20 cities with the largest positive gaps, 19 are in developed countries, with Da Nang, Vietnam, as the sole exception from a less-developed country.

3.4. Regression Results

We now assess whether the patterns documented in Section 3.2 (distance to the city center) and Section 3.3 (hills and rivers) hold 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 city-year observation,[6] 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-year fixed effects ν_c, and the error term $\epsilon_{j,c}$ captures idiosyncratic

[6]Recall that we have four cities with multiple years of observations from our JICA travel survey.

neighborhood-level variation. We weight observations by the inverse of the fraction of residents in each neighborhood, such that the regression assigns equal weight to each city. We cluster standard errors at the city-year 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. When disaggregated between hilly and river areas, the coefficients for developed cities are larger for hilly areas (14.6 percentile points) than for river areas (2.9 percentile points), with both estimates statistically significant. In contrast, for less-developed cities, the coefficients are similar and negative (−7.1 and −7.3 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 0.16 log points higher in developed cities, and 0.28 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.

Robustness In Table 3, we show that the main patterns remain robust across a variety of alternative specifications. Row (1) replicates the baseline estimates from the bottom panel of Column (2) in Table 1, confirming statistically significant differences in income premiums associated with suburban, hilly, and river neighborhoods between less developed and developed cities.

Rows (2) and (3) exclude neighborhoods located more than 15 km and 20 km from the city center, respectively. While the estimated income percentile differentials become slightly smaller, they remain statistically significant, indicating that our results are not driven by the definition of city boundaries or differences in city size across development levels.

Rows (4) and (5) exclude cities in the USA and Brazil, respectively, which represent a large

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		
$\text{Developed}_c \times \text{Suburban}_{j,c}$	10.2***	0.16***
	(1.9)	(0.03)
$\text{Less Developed}_c \times \text{Suburban}_{j,c}$	-13.8***	-0.28***
	(1.6)	(0.04)
$\text{Developed}_c \times \text{Hilly}_{j,c}$	14.6***	0.20***
	(3.1)	(0.04)
$\text{Less Developed}_c \times \text{Hilly}_{j,c}$	-7.1***	-0.17***
	(1.7)	(0.04)
$\text{Developed}_c \times \text{River}_{j,c}$	2.9**	0.06***
	(1.1)	(0.02)
$\text{Less Developed}_c \times \text{River}_{j,c}$	-7.3***	-0.07**
	(1.8)	(0.03)
Difference: Less Developed$_c$ vs Developed$_c$		
$\text{Suburban}_{j,c}$	-24.0***	-0.44***
	(2.5)	(0.04)
$\text{Hilly}_{j,c}$	-21.7***	-0.36***
	(3.5)	(0.06)
$\text{River}_{j,c}$	-10.2***	-0.13***
	(2.1)	(0.03)
Observations	145,377	145,357
Unique City-Years	132	132
City-Year FE	✓	✓
Weight by neighborhood pop within city	✓	✓

Note: Top panel reports the results of the regression (1). Bottom panel reports the results of the regression where we replace $X_{j,c} \times Developed_c$ with $X_{j,c}$ in the regression (1) to assess the differences in the coefficients between developed and less-developed cities. Unit of observation is a neighborhood. We weight observations by the inverse of the fraction of residents in each neighborhood, such that the regression assigns equal weight to each city. Standard errors are clustered at city-year level. ***, ** and * indicate statistical significance at the 1-percent, 5-percent and 10-percent levels.

share of our sample (48 and 24 cities, respectively). Excluding the USA – where the income premiums in suburban, hilly, and river areas are the greatest – reduces the magnitude of the differences somewhat, but the estimates remain statistically significant. Similarly, the

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	-24.0 (2.5)***	-21.7 (3.5)***	-10.2 (2.1)***
2	Exclude neighborhoods $\geq$ 15km of center	-18.7 (2.5)***	-19.0 (4.1)***	-8.7 (2.3)***
3	Exclude neighborhoods $\geq$ 20km of center	-22.6 (2.5)***	-20.3 (3.9)***	-9.0 (2.3)***
4	Exclude USA cities	-12.2 (2.9)***	-15.4 (4.9)***	-7.3 (2.8)***
5	Exclude Brazil cities	-22.9 (3.2)***	-17.0 (5.8)***	-10.4 (2.4)***
6	New World cities	-31.4 (2.8)***	-29.9 (3.4)***	-12.4 (2.4)***
7	Old World cities	-10.6 (3.7)***	-6.0 (6.2)	-7.5 (3.1)**
8	Control for neighborhood log area	-21.7 (2.4)***	-22.2 (3.5)***	-6.4 (1.8)***
9	Control for neighborhood quadrant to center	-24.3 (2.5)***	-20.2 (3.6)***	-10.1 (2.1)***
10	Control for city population	-23.3 (2.7)***	-21.6 (3.5)***	-9.9 (2.4)***
11	Control for city area	-23.5 (2.9)***	-18.8 (3.9)***	-10.4 (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 United States 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-year level. ***, ** and * indicate statistical significance at the 1-percent, 5-percent and 10-percent levels. Appendix Table B.5 reports additional robustness, such as using alternative proxies for distance to city center or controlling for city-level characteristics, such as ethnic diversity.*

exclusion of Brazil slightly attenuates the results, while the results on the differences between less developed and developed cities hold robustly.

Rows (6) and (7) restrict the sample to "New World" cities (North and Latin America) and "Old World" cities (all others), respectively. 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. This suggests that our findings are not merely driven by the New World vs. Old World distinction, which has been noted as a key factor shaping subnational economic geography (Henderson et al., 2018).

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 et al., 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.

4. Quantitative Urban Model

In this section, we develop a quantitative framework that we use to assess various potential mechanisms in explaining the gap of spatial income distribution between cities in less-developed and developed countries. Our analysis is based on a quantitative urban model (Redding and Rossi-Hansberg, 2017; Redding, 2023), which explicitly model each city as a collection of neighborhoods with flexible heterogeneity in geographic features. This approach enables us to fit our model to neighborhood-level data for each city, disentangle the key forces shaping spatial income patterns, and compare systematic differences between less-developed and developed cities beyond city-specific idiosyncraticies.

Our baseline model features four basic potential explanations of the gap. The first is non-homothetic preference over amenities, such that richer households care more about neighborhood amenities like those available in leafy suburbs far from the city center, or the views from hilly areas or of water. The second is the comparatively worse transportation infrastructure and higher commuting costs in less developed countries. The third is more centralized jobs in developing countries, which we allow as a possibility in the model by having a steeper productivity decline in distance from city center. The fourth is the possibility that amenity gradients decline faster with distance from city center, or are relatively worse in hills and rivers, in less developed cities. We relegate another potential explanation based on the heterogeneity in commuting costs across income groups to Section 6.3.

4.1. Environment

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.

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).

4.2. Households' Preferences and Residential Location Decisions

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, 2022; 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 United States (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:

$$1 = \left(\frac{y}{U_j}\right)^{\frac{\sigma-1}{\sigma}} + \chi^{\frac{1}{\sigma}} \left(\frac{h}{U_j^{\varepsilon}}\right)^{\frac{\sigma-1}{\sigma}} \tag{2}$$

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 (2) 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, 2022), 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 cities.[7]

[7] Alternative classes of non-homothetic preferences commonly used in studies of individual cities or countries, such as Stone-Geary preferences (Tsivanidis, 2023) 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.

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

$$\log\left(U_j^{\rho}+B_j^{\rho}\right)^{\frac{1}{\rho}}+\epsilon_j(\omega) \tag{3}$$

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 (2023); Finlay and Williams (2022); 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:

$$\begin{aligned} U_j(s(\omega)) \equiv \arg\max_{\{h,y\}} \quad & U_j \\ \text{s.t.} \quad & P_j h + y \le \overline{w}_j s(\omega) \quad \text{and Equation (2)} \end{aligned} \tag{4}$$

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 4.4.

Anticipating this decision, household ω chooses the residential location that maximizes the utility (3)

$$j(\omega)\equiv \arg\max_j V_j(s(\omega))+\epsilon_j(\omega), \qquad V_j(s)\equiv \log\left(U_j(s)^{\rho}+B_j^{\rho}\right)^{\frac{1}{\rho}} \tag{5}$$

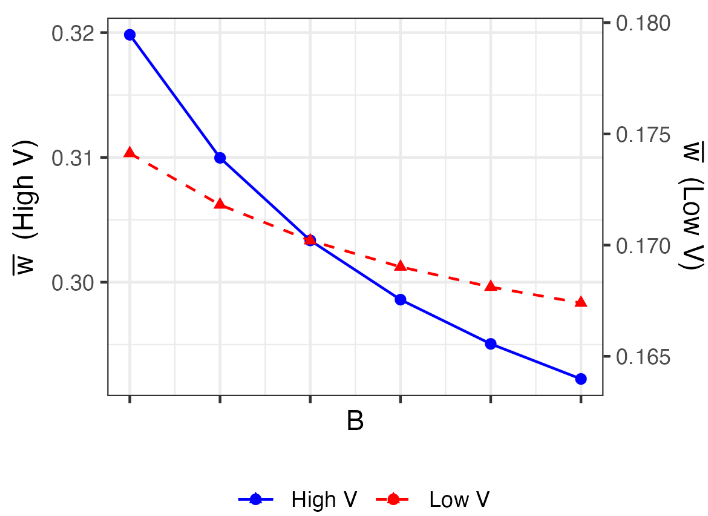
Following the literature, we assume that $\epsilon_j(\omega)$ is independently drawn from the Gumbel distribution with scale parameter ν. Then, the probability that households with earning potential s reside in location j is given by

$$\pi_j(s)=\frac{\exp\left(\nu V_j(s)\right)}{\sum_{\ell}\exp\left(\nu V_{\ell}(s)\right)} \tag{6}$$

4.3. 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 6, 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 6: 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.[8]

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, 2022).

4.4. 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. 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

$$n(v) = \arg\max_n \tau_{jn}^{-1} w_n \tilde{\epsilon}_n(v). \tag{7}$$

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

[8] At first glance, it may seem puzzling that the constant-elasticity-of-substitution specification with B_j (Equation 3) 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.

n is given by

$$\lambda_{jn} = \frac{\left(\tau_{jn}^{-1} w_n\right)^{\theta}}{\sum_{\ell}\left(\tau_{j\ell}^{-1} w_{\ell}\right)^{\theta}}. \tag{8}$$

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

$$\overline{w}_j = \varrho \sum_{\ell}\left(\tau_{j\ell}^{-1} w_{\ell}\right)^{\theta}, \tag{9}$$

where $\varrho \equiv \Gamma\left(\frac{\theta-1}{\theta}\right)$, where $\Gamma(\cdot)$ is the Gamma function. Equation (9) indicates that the wage rates per efficiency unit of labor for residents in neighborhood j is proportional to the geometric average of the wage rates at various work destinations weighted by the commuting cost, often called "commuting access" in the literature (e.g., Tsivanidis, 2023).[9]

4.5. Production, Market Clearing, and Equilibrium

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

$$w_n = A_n. \tag{10}$$

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 (Saiz, 2010). We assume that the inverse supply function of housing is given by

$$P_j = \frac{1}{S_j} H_j^{\mu}, \tag{11}$$

where H_j is the aggregate supply of housing.

The market clearing of housing in location j is given by

$$H_j = \int_s h(s)\pi_j(s) dG(s), \tag{12}$$

where $G(\cdot)$ is the cumulative distribution function of earning potential $s(\omega)$ across households.

The equilibrium is defined by households' consumption $\{h(s), y(s)\}$, residential choice

[9] We assume a continuum of households, 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 4.2.

probabilities $\{\pi_j(s)\}$, labor supply probabilities $\{\lambda_{jn}\}$, wages $\{w_j\}$, and house prices $\{P_j\}$, that satisfy households' optimal consumption decision (4), residential location decision (5) and (6), labor supply decision (7) and (8), final goods producers' optimality condition (10), and housing supply and market clearing conditions (11) and (12).

The average residential income in neighborhood j is given by

$$I_j = \overline{w}_j \frac{\int_s s\pi_j(s)dG(s)}{\int_s \pi_j(s)dG(s)}, \tag{13}$$

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 (9). 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, as discussed in Section 4.3.

5. Measuring Commuting Costs and Access to Jobs for Each City

We now combine the framework developed in Section 4 with our data to quantify the contributions of different mechanisms to the observed gap in spatial income distributions between developed and less-developed cities. In this section, we estimate city-level commuting costs using the model-implied commuting gravity equations. In the next section, we employ the full general equilibrium model to conduct counterfactual analyses and assess the underlying drivers of the spatial income distribution gap.

5.1. Estimating Commuting Gravity Equations for Each City

We use our model's commuting (labor supply) decisions in Section 4.4 to quantify the commuting costs for each city. Adding the city subscripts c for all of our model variables, our model predicts the probability of commuting by residents (household members) in residential location j to workplace location n by

$$\lambda_{jn,c} = \frac{\left(\tau_{jn,c}^{-1} w_{n,c}\right)^{\theta}}{\sum_{\ell}\left(\tau_{j\ell,c}^{-1} w_{\ell,c}\right)^{\theta}}, \tag{14}$$

where c indicates a city. We introduce a parametric assumption that the commuting cost $\tau_{jn,c}$ is a power function of road distance between neighborhoods j to n such that

$$\log \tau_{jn,c} = \kappa_c \text{RoadDistance}_{jn,c}, \tag{15}$$

where we allow κ_c to vary by city, and $\epsilon_{jn,c}$ is additional idiosyncratic travel costs for a pair of neighborhoods that are not captured by road distance.

Combining the two equations and taking logs,

$$\log \lambda_{jn,c} = \theta \log w_{n,c} - \kappa_c \theta \text{RoadDistance}_{jn,c} - \log \sum_{\ell}\left(\tau_{j\ell,c}^{-1} w_{\ell,c}\right)^{\theta}. \tag{16}$$

We estimate an empirical analog of this equation using a Pseudo-Poisson Maximum Likelihood (PPML) estimator to handle the presence of many pairs of zero commuters. Specifically, we estimate the following two-way fixed model

$$\log \mathbb{E}\left[\lambda_{jn,c}\right] = \psi_{n,c} - \tilde{\kappa}_c \text{RoadDistance}_{jn,c} + \eta_{j,c}, \tag{17}$$

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

The estimates from Equation (17) yield three key insights into the structure of commuting costs within each city. First, the estimated commuting semi-elasticity with respect to road distance, $\tilde{\kappa}_c$, captures the intensity of commuting costs and allows for comparisons across cities. Second, given an assumed value of θ, the estimated workplace fixed effects $\psi_{n,c}$ can be interpreted as proportional to the wage rates (or productivity) at each workplace location, since $w_{n,c} = A_{n,c} \propto \exp\left(\psi_{n,c}/\theta\right)$ (Kreindler and Miyauchi, 2023). Third, the origin fixed effects $\eta_{n,c}$ are proportional to wage rates at residential locations, following $\overline{w}_{j,c} \propto \exp\left(-\eta_{n,c}/\theta\right)$ (Equation 9).[10] This last component is particularly important, as it directly enters the determination of equilibrium residential income via Equation (13).

[10]Under the PPML specification, 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: $\sum_{\ell} \exp\left(\hat{\tilde{\kappa}}_c \text{RoadDistance}_{j\ell,c} + \hat{\psi}_{\ell,c}\right)$ (Fally, 2015).

5.2. Estimation Results

In Panel (a) of Figure 7, we present the point estimates of $\tilde{\kappa}_c \equiv \kappa_c \theta$ from estimating equation (17) for each city. We estimate these equations for the U.S. cities, Tokyo, and less-developed cities from the JICA travel survey data. We exclude cities from the UK, France, Spain, and Brazil, where we do not have bilateral commuting flow data.

The figure highlights the striking heterogeneity in commuting costs across cities. For U.S. cities, the mean of the estimated semi-elasticities is approximately 0.07, implying that a one-kilometer increase in commuting distance leads to a 7% reduction in commuting flows, conditional on origin and workplace fixed effects. In Tokyo, the corresponding estimate is approximately 0.11. This difference between the U.S. and Japan (Tokyo) may reflect variation in typical commuting modes: in the U.S., private cars are the predominant mode of transport, whereas in Tokyo, public transit is more common.

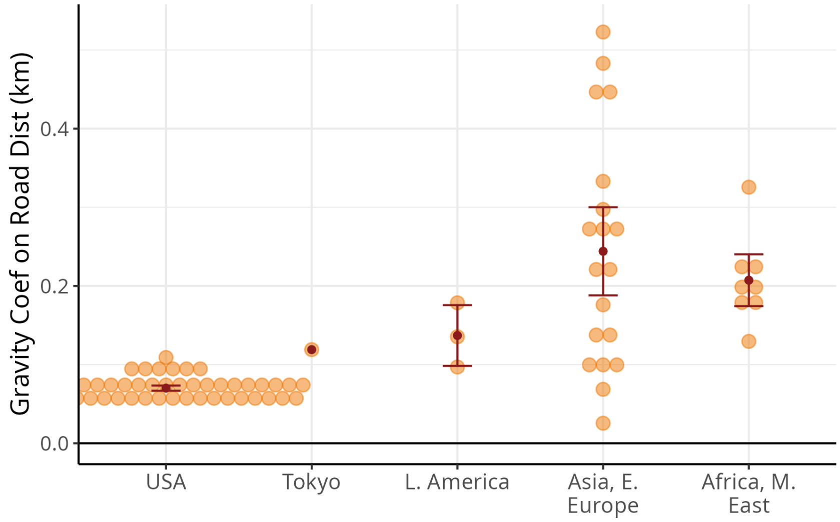
In contrast to these developed cities, the semi-elasticity of commuting is significantly more negative and larger in magnitude in less-developed settings. For Latin American cities, the mean estimate is 0.14; for cities in Asia and Eastern Europe, it rises to 0.24; and for cities in Africa and the Middle East, it is 0.18. These higher values may reflect limited access to private vehicles, less developed road infrastructure, and the absence of efficient public transit systems.[11]

Panel (b) of Figure 7 plots the estimated commuting semi-elasticities against the negative speed index from Akbar et al. (2023b), which measures log-point differences in average road speed across cities. We observe a statistically significant positive relationship: in cities with slower average road speeds (higher values on the horizontal axis), commuting semi-elasticities tend to be higher (higher values on the vertical axis). The estimated slope is 1.47 and significantly greater than one, suggesting that commuting frictions increase more than proportionally as road speeds decline. This pattern is consistent with the interpretation that in faster (typically richer) 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. Furthermore, the relationship exhibits considerable dispersion, suggesting that other factors may also influence commuting behavior, potentially including road safety or cultural norms.

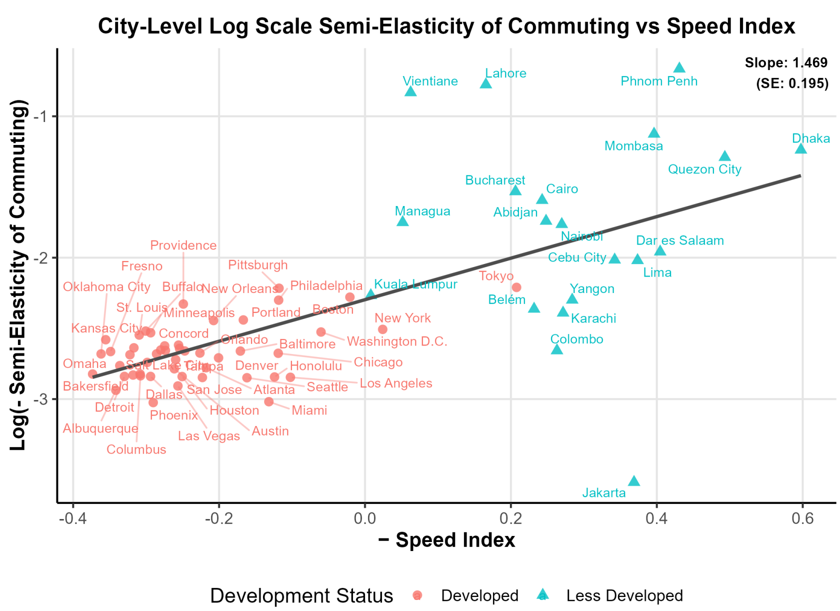
We next focus on our estimates of productivity at workplaces $A_{n,c} \propto \exp\left(\psi_{n,c}/\theta\right)$ and wage

[11] Our estimated semi-elasticities align with existing city-specific estimates reported by Tsivanidis (2023); Zárate (2024); Khanna et al. (2023), who find values in the range of 0.02 to 0.05 with respect to *minutes* instead of *kilometers*, based on assumed mode-specific travel speeds for each city.

Figure 7: Estimated Semi-Elasticity of Commuting to Road Distance ($\tilde{\kappa}_c$)



(a) By Continents



(b) Relationship with Speed on Roads

Note: Results of the estimated coefficients of road distance using equation (17) for each city. In Panel (a), each dot corresponds to an estimate from each city. We also report the group means of the point estimates along with 95 percent confidence intervals of these point estimates across cities. In Panel (b), the horizontal axis displays the negative speed index from Akbar et al. (2023b), which measures the log-point difference in average road speed in each city. See Appendix Table C.1 for the estimates for each city in less developed countries.

rates at residential location $\overline{w}_{j,c} \propto \exp\left(\eta_{n,c}/\theta\right)$. To estimate workplace productivity, we adjust the workplace fixed effects $\psi_{n,c}$ by subtracting the log area size, addressing the mechanical one-to-one relationship between area and fixed effects noted in Kreindler and Miyauchi (2023).[12]

In Table 4, we present how these variables relate to the geographic features highlighted in Section 3. Specifically, we run the same regression (1), where we regress these estimates on the suburban, hilly, and river dummies interacted with developed and less-developed city dummies. As before, top panel reports the results of regression (1), and the bottom panel reports the differences in the coefficients between less developed versus developed cities. We set $\theta = 5$, consistent with median estimates reported in the literature for both developed and less-developed cities.[13]

In Column (1), we find that workplace productivity is significantly lower in suburban, hilly, and river areas in both developed and less-developed cities. This is consistent with the interpretation that productivity tends to be higher near city centers. However, the magnitudes of these effects differ notably across development levels. In particular, the productivity disadvantage in suburban and river areas is substantially larger in less-developed cities compared to developed ones. Interestingly, the productivity disadvantage in hills are smaller in less-developed cities compared to developed ones. This pattern perhaps reflect the fact that in developed cities, hilly areas specialize in residential purposes, and residents there simply work to some other neighborhoods.

In Column (2), we turn to the wage rates *at residential locations*, which reflect commuting access. These rates also tend to be lower in suburban, hilly, and river areas in both developed and less-developed cities. This pattern supports the interpretation that areas farther from city centers have lower access to jobs, and that commuting costs transmit these spatial differences into residential wage rates. Once again, the magnitude of these effects is more pronounced in less-developed cities. In the bottom panel, the wage penalty in suburban areas is 0.10 log points larger in less-developed cities, while the penalty around rivers is 0.05 log points larger. For hilly areas, the estimated penalty is slightly positive at 0.02 log points, though statistically insignificant.

In Figure 8, we examine the relationship between observed residential income and the estimated commuting access. Specifically, for both less-developed and developed cities, we compute the average commuting access (the negative estimated destination fixed effects from

[12]This adjustment can be microfounded by assuming that workers draw idiosyncratic productivity shocks per unit of geographic area (see Kreindler and Miyauchi (2023) for details).

[13]The estimates from the prior research range from 2.2 to 8.3 (e.g., Ahlfeldt, Redding, Sturm, and Wolf, 2015; Kreindler and Miyauchi, 2023; Severen, 2023; Tsivanidis, 2023).

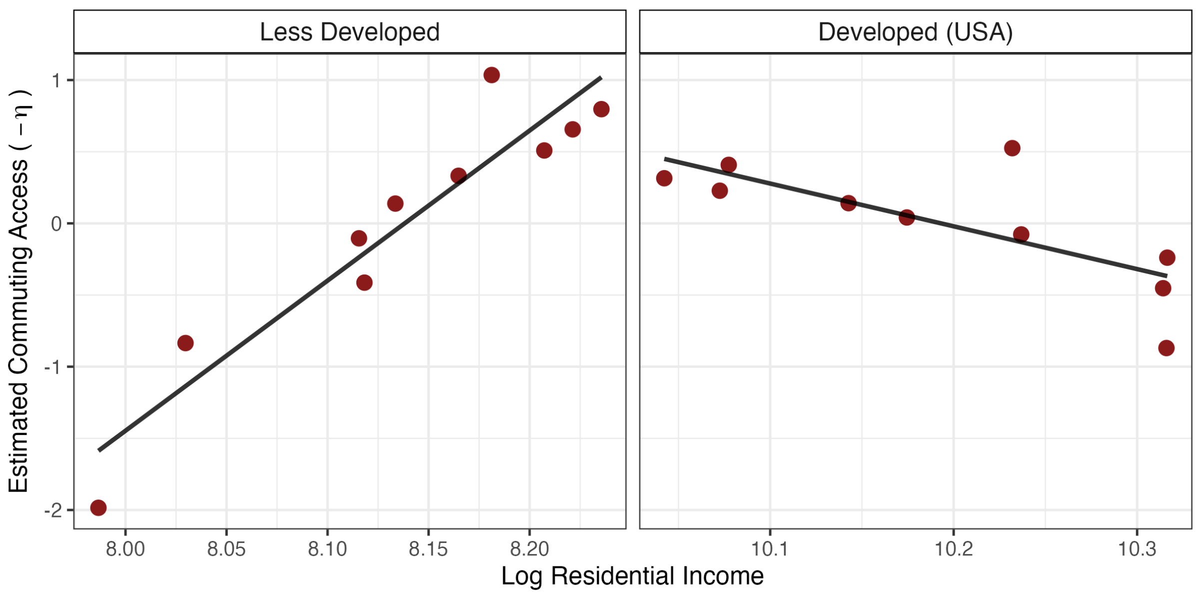
Table 4: Regression Results of Estimated A_j and $\overline{w}_j$ on Suburban, Hilly, and River Dummies

Dependent Variables:	$\log A_{j,c}(= [\psi_{j,c} - \log Area_{j,c}]/\theta)$	$\log \overline{w}_{j,c}(= -\eta_{j,c}/\theta)$
Model:	(1)	(2)
Variables		
$\text{Developed}_c \times \text{Suburban}_{j,c}$	-0.11***	-0.09***
	(0.01)	(0.008)
$\text{Less Developed}_c \times \text{Suburban}_{j,c}$	-0.27***	-0.19***
	(0.02)	(0.02)
$\text{Developed}_c \times \text{Hilly}_{j,c}$	-0.26***	-0.03***
	(0.03)	(0.01)
$\text{Less Developed}_c \times \text{Hilly}_{j,c}$	-0.14***	-0.05*
	(0.03)	(0.03)
$\text{Developed}_c \times \text{River}_{j,c}$	-0.05***	-0.02***
	(0.009)	(0.004)
$\text{Less Developed}_c \times \text{River}_{j,c}$	-0.16***	-0.07***
	(0.02)	(0.02)
Difference: Less Developed$_c$ vs Developed$_c$		
$\text{Suburban}_{j,c}$	-0.16***	-0.10***
	(0.02)	(0.02)
$\text{Hilly}_{j,c}$	0.12***	-0.02
	(0.04)	(0.03)
$\text{River}_{j,c}$	-0.11***	-0.05***
	(0.02)	(0.02)
Observations	33,279	31,253
City-Year FE	✓	✓
Unique City-Years	79	71
Weight neghborhoods equally within city	✓	✓
USA, Tokyo, JICA Cities	✓	✓

Note: Top panel reports results from regression (1) using an alternative dependent variable. Bottom panel reports the results of the regression where we replace $X_{n,c} \times Developed_c$ with $X_{n,c}$ in the regression (17) to assess the differences in the coefficients between developed and less-developed cities. In Column (1), the dependent variable is the estimated destination fixed effect from the commuting gravity equation (17), net of log area and scaled by $\theta = 5$. Column (2) uses the estimated origin fixed effects from the same equation as the dependent variable.

the commuting gravity equation 17), and plot them against observed residential income. Given that most of our developed-country commuting data comes from the United States, we restrict the developed group to U.S. cities and exclude Tokyo, the only other developed-country city with such data.

Figure 8: Estimated Commuting Access and Residential Income



Note: For both less-developed and developed (U.S.) cities, neighborhoods are divided into city-specific income deciles. For each decile, we compute the mean of the negative estimated destination fixed effects from the commuting gravity equation (17). Each dot represents the average value for a given income decile, aggregated across all cities within each group (left: less-developed cities; right: developed (U.S.) cities).

The figure reveals a striking contrast between less-developed and developed cities. In less-developed cities (left panel), there is a clear, monotonic positive relationship: neighborhoods with higher residential income also enjoy better commuting access. In contrast, U.S. cities (right panel) exhibit a monotonic negative relationship: higher-income neighborhoods are associated with lower commuting access. This pattern suggests that, in the U.S., households with higher earning potential tend to sort into neighborhoods with lower commuting access – such as leafy suburbs or areas with scenic amenities like hills or waterfronts. In less-developed cities, similar sorting may occur but appears too weak to overturn the strong positive relationship between commuting access and residential income. Therefore, neighborhoods with lower commuting access – in particular, suburban, hilly, or river areas, which are less well-connected by transportation infrastructure – tend to have lower residential income.[14] In

[14] This observation echoes the findings of Kreindler and Miyauchi (2023), who show that commuting access is a strong predictor of income in Dhaka, Bangladesh, and Colombo, Sri Lanka – two less-developed cities included in our dataset. In contrast, our results suggest that this relationship may not hold in developed-country cities.

the next section, we quantify how these patterns jointly determine the equilibrium residential location choice and spatial income distribution.

6. Quantifying the Gap in Equilibrium Spatial Income Distribution

In this section, we use the full general equilibrium model to conduct counterfactual simulations that identify the key drivers of the spatial income distribution gap between developed and less-developed cities. We begin by calibrating the model to U.S. cities as a benchmark for developed economies, then implement a series of counterfactual exercises to disentangle the mechanisms underlying the observed disparities. We also discuss the distributional impacts of a city's overall productivity growth across households.

6.1. Calibrating the Model to U.S. cities

We begin by introducing additional parametric assumptions for the amenity term $B_{j,c}$, modeling it as a function of observable geographic features:

$$\log B_{j,c} = \beta_0 + \beta_1 \text{Suburban}_{j,c} + \beta_2 \text{Hills}_{j,c} + \beta_3 \text{Rivers}_{j,c}, \tag{18}$$

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). We assume $\{\beta_0, \beta_1, \beta_2, \beta_3\}$ as common preference parameters across across all cities.[15]

In the previous section, we calibrated the shape parameter for productivity shocks θ as well as productivity $\{A_{j,c}\}_j$ and commuting costs $\{\tau_{jn,c}\}_{j,n}$ for each city and neighborhood pairs.[16] We now turn to the calibration of the remaining parameters, as summarized in Table 5.

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 (2022), 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

[15] 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. While this choice implies that our model does not exactly match the observed residential income distribution, we estimate the key parameters to rationalize the key moments, such as the relative incomes in suburban, hilly, and river areas.

[16] The procedure in the previous section only reveals the relative $\{A_{j,c}\}_j$ within the city, but not its scale. We set the scale to replicate the observed mean income of U.S. cities.

Table 5: Calibrated Parameters Targeting U.S. cities

	Value	Description	Source
θ	5	Dispersion of idiosyncratic productivity shock	Literature
σ	0.52	Elasticity of substitution for housing	Finlay and Williams (2022)
ε	0.36	Elasticity of nonhomotheticity in housing	Finlay and Williams (2022)
$G(\cdot)$	{3,1} wp 0.5	Earning potential distribution	Variance of income distribution
χ	0.063	Preference shifter for housing	GMM
ν	11.8	Residential location choice elasticity	GMM
ρ	-0.43	Elasticity of nonhomotheticity in amenity	GMM
β_0	9.25	Amenity level	GMM
β_1	0.55	Amenity for suburban areas	GMM
β_2	0.85	Amenity for hilly areas	GMM
β_3	0.18	Amenity for river areas	GMM
μ	0.5	Inverse housing supply elasticity	Literature

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

patterns observed in the United States 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 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. Specifically, using observed median rents $\{P_{j,c}\}$ from ACS and the estimated wage rates at residential locations $\{\overline{w}_{j,c}\}$ in Section 5, 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, given below:

1. Log average residential income in location j, $\log I_j$ (Equation 13), as well as its interaction with suburban, hilly, and river dummies.

2. The interaction of log average residential income and the commuting access $\log \overline{w}_j \propto \frac{1}{\theta} \log \sum_{\ell} \left(\tau_{j\ell,c}^{-1} w_{\ell,c} \right)^{\theta}$, as estimated in Section 5.1. To address concerns that the placement of roads may be endogenously correlated with unobserved residential amenities, we modify our measure of commuting access by replacing actual road distances for the estimation of Equation (17) with bilateral straight-line distances.

This alternative specification helps mitigate potential endogeneity arising from the endogenous assignment of road infrastructure.

3. The interaction of the moments described in point 2 above and the dummies of hills and rivers.

4. Average housing expenditure share of residents in location j.

5. Implied residential location choice elasticity with respect to wage rates at residences (commuting access). 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, 2023; Severen, 2023).

Finally, we choose the value that minimizes the GMM objective function:

$$\hat{\Theta} = \min_{\Theta} \left\{ \left(\frac{1}{N} \sum_{n,c} g_{j,c}(\Theta) \right)' \mathbf{W} \left(\frac{1}{N} \sum_{n,c} g_{j,c}(\Theta) \right) \right\}, \tag{19}$$

where $\mathbf{W}$ is the weighting matrix.

We now describe how each parameter $\Theta \equiv \{\beta_0, \beta_1, \beta_2, \beta_3, \rho, \nu, \chi\}$ is identified from these moments. 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 $\{\beta_0, \beta_1, \beta_2, \beta_3\}$, as higher values of these parameters imply that households with higher earning potential s are more likely to sort into suburban areas, hills, and river-adjacent neighborhoods, holding other factors constant (given $\rho < 0$). The second moments are informative about the value for ρ, since a more negative ρ implies that households with lower earning potential are disproportionately more responsive to commuting access. The third set of moments are informative about the value of β_0, as its level determines how the relationship between commuting access and earning potential varies systematically across different amenity neighborhoods. 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 6 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 as estimated in Section 5. 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 5, except we replace actual road distances for the estimation of Equation (17) with bilateral straight-line distances, consistent with the moment condition used above.

Table 6: 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			0.27	0.27	0.27	0.27
	(0.01)	(0.03)			(0.01)	(0.04)	(0.01)	(0.04)
Hilly$_{j,c}$	0.29	0.29			0.30	0.30	0.31	0.31
	(0.01)	(0.05)			(0.01)	(0.06)	(0.01)	(0.06)
River$_{j,c}$	0.06	0.07			0.07	0.07	0.07	0.08
	(0.004)	(0.02)			(0.004)	(0.02)	(0.003)	(0.02)
$\overline{w}_{j,c}$			−0.79	−0.79	0.44	0.44	0.31	0.36
			(0.11)	(0.23)	(0.08)	(0.29)	(0.08)	(0.28)
$\overline{w}_{j,c} \times \text{Hilly}_{j,c}$							0.71	0.49
							(0.12)	(0.46)
$\overline{w}_{j,c} \times \text{River}_{j,c}$							0.29	0.16
							(0.05)	(0.24)
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 5, using the value of* $\theta = 5$. *We instrument* $\overline{w}_{j,c}$, $\overline{w}_{j,c} \times \mathit{Hilly}_{j,c}$ *and* $\overline{w}_{j,c} \times \mathit{River}_{j,c}$ *by* $\tilde{\overline{w}}_{j,c}$, $\tilde{\overline{w}}_{j,c} \times \mathit{Hilly}_{j,c}$, *and* $\tilde{\overline{w}}_{j,c} \times \mathit{River}_{j,c}$, *where* $\tilde{\overline{w}}_{j,c}$ *is constructed analogously as in Section 5, except we replace actual road distances for the estimation of Equation (17) with bilateral straight-line distances. Standard errors are clusterred at the city level.*

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 commuting access, consistent with the pattern illustrated in Figure 8. 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 commuting access, both in the model and data. Finally, Columns (7) and (8) show that the interaction between commuting 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. We therefore use this estimated model as our baseline for U.S. cities in the subsequent counterfactual simulations.

Finally, we calibrate the 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 $P_{j,c}$.

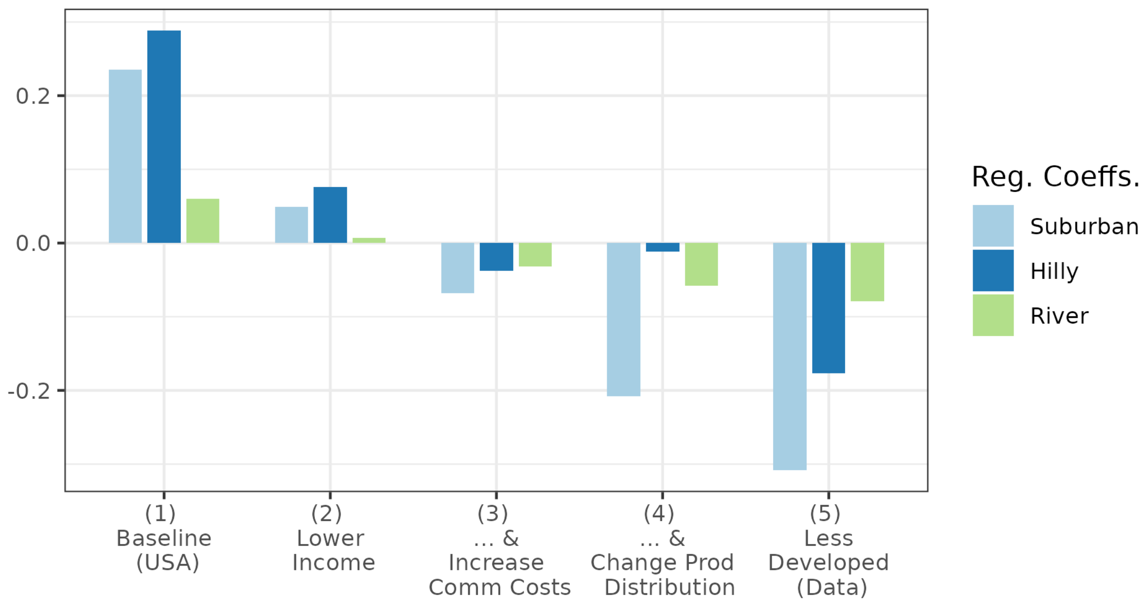
6.2. Counterfactuals

We now use our calibrated model for U.S. cities to conduct a series of counterfactual simulations. Specifically, 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 ($\tilde{\kappa}_c$) to the levels estimated for less-developed cities in Section 5; 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, as reported in Table 4.

Figure 9 presents the results. In each column, we report the results of the estimated coefficients on suburban, hilly, and river dummies on residential income using regression specification (1). Column (1) reports the patterns using our calibrated model to U.S. cities, as reported in Column (1) of Table 6. Columns (2)-(4) report the regressions using U.S. cities under counterfactual equilibrium under alternative scenarios, as we further decsribe below. Column (5) report the regression coefficients for less developed cities using our data, as reported in Table 2. Consistent with the findings so far, Column (1) exhibit positive income premiums in suburban, hills, and rivers for the U.S. cities, and Column (5) exhibit negative income premiums in those areas. We explore whether and what type of counterfactual close this observed spatial income gap.

In Column (2), we present results from a counterfactual in which average productivity

Figure 9: Spatial Residential Income Distribution of U.S. Cities under Counterfactual Scenarios



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 6. 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 $\tilde{\kappa}_c$ by 0.14, reflecting the average gap between U.S. and less-developed cities as shown in Figure 7. 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, as documented in Table 4. Finally, Column (5) displays the corresponding estimates for less-developed cities based on observed data, as reported in Table 2.

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.[17] Under this scenario, the regression 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

[17] Due to endogenous adjustments in residential and commuting 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.

sorting into amenity-rich areas.

In Column (3), we run a counterfactual to additionally increase the commuting semi-elasticity $\tilde{\kappa}_c$ by 0.14, matching the average difference in estimated $\tilde{\kappa}_c$ between U.S. and less-developed cities (Figure 7). 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 commuting access in those areas.[18]

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, as documented in Table 4. We find that the negative income premiums in suburban and river areas become more pronounced. In contrast, the negative income premium in hilly areas diminishes somewhat, reflecting the smaller productivity penalty associated with hilly neighborhoods in less-developed cities, as shown in Table 4.

Figure 9 shows that the differences in overall income levels, commuting costs, and to a lesser extent, spatial productivity distributions, explain the major gap in observed spatial income distribution between developed and less-developed cities. However, there remains a residual gap compared to the patterns observed in less-developed cities. What accounts for this remaining discrepancy? One potential explanation is heterogeneity in commuting costs across income groups (Glaeser et al., 2008; Su, 2022). However, as discussed in Section 6.3, this mechanism appears unlikely to play a major role. Another plausible explanation is lower amenity valuations in suburban, hilly, or river-adjacent areas in less-developed cities. 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 (McCulloch et al., 2025).[19]

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 7 shows that reducing these coefficients from the baseline values of $\{0.55, 0.85, 0.18\}$ to lower but positive values of $\{0.29, 0.36, 0.12\}$ would bridge the observed gap. 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. Importantly, we do not need to

[18] Appendix Figure C.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 the largest effects, followed by increasing commuting costs, then changing productivity distribution.

[19] Another part of the gap between Columns (5) and (6) 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.

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.

Table 7: 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)	0.55	0.85	0.18
Estimates to Fully Rationalize Less Developed Cities	0.29	0.36	0.12

Note: First row reports the estimated $\{\beta_1, \beta_2, \beta_3\}$ using U.S. cities as reported in Table 5. 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.

6.3. Heterogeneity of Commuting Costs 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 4 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:

$$\lambda_{jn,c}(s) = \frac{\left(\tau_{jn,c}(s)^{-1} w_{n,c}(s)\right)^{\theta}}{\sum_{\ell}\left(\tau_{j\ell,c}(s)^{-1} w_{\ell,c}(s)\right)^{\theta}}. \tag{20}$$

We also follow Section 5 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 (16) holds separately for each earning potential s:

$$\log \lambda_{jn,c}(s) = \theta \log w_{n,c}(s) - \kappa_c(s)\theta \times \text{RoadDistance}_{jn,c} - \log \sum_{\ell}\left(\tau_{j\ell,c}(s)^{-1} w_{\ell,c}(s)\right)^{\theta}. \tag{21}$$

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 (17):

$$\log \mathbb{E}\left[\lambda_{jn,c}(s)\right] = \psi_{n,c} - \tilde{\kappa}_c(s)\text{RoadDistance}_{jn,c} + \eta_{j,c}(s), \tag{22}$$

where $\psi_{n,c}(s) \equiv \theta \log w_{n,c}(s)$ is the workplace fixed effects, $\tilde{\kappa}_c(s) \equiv \kappa_c(s)\theta$, 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.

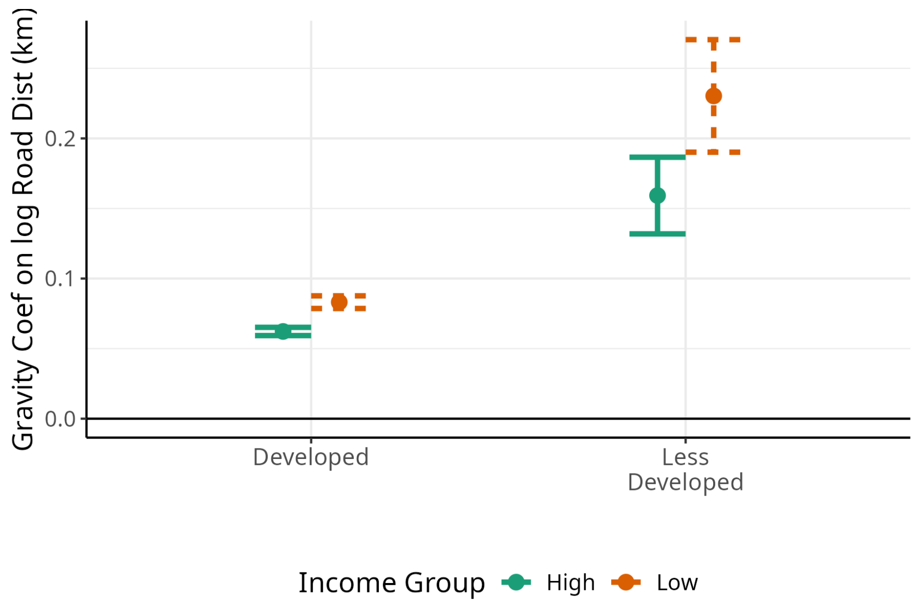
A challenge in estimating this equation is that we do not directly observe earning potentials s in our data. While some travel survey data provide information about variables such as education, it is not comprehensive, and it is also difficult to create a proxy that is consistent across countries. To 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 10 presents the estimated commuting semi-elasticities based on Equation (22). 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), and the right column shows those estimates in less-developed cities. In both groups of cities, low-income households exhibit higher commuting semi-elasticities than high-income households. This pattern is consistent with the notion that higher-income households have better access to more efficient modes of transportation, such as private vehicles (Glaeser et al., 2008; Tsivanidis, 2023). At the same time, the differences in semi-elasticities between developed and less-developed cities are significantly larger than within-city across-income-group variations.[20]

What are the implications of these heterogeneity in commuting costs for the spatial distribution of income? To explore this question, we conduct the following counterfactual exercise. First, we calibrate our extended model to U.S. cities using the same parameters as in our baseline specification (Table 5). To incorporate heterogeneity in commuting access, we use the estimated wage rates at residential locations for high- and low-earning-potential groups, $\overline{w}_{j,c}(s)$ for $s \in \{H, L\}$, obtained from the gravity equation (22). Second, we simulate a scenario in which this heterogeneity is eliminated by setting $\overline{w}_{j,c}(L) = \overline{w}_{j,c}(H)$. That is, we assume that low- and high-earning-potential groups face the same commuting access. We then assess how

[20] Appendix Table C.2 reports the patterns of estimated destination and origin fixed effects.

Figure 10: Semi-Elasticity of Commuting by Income Groups



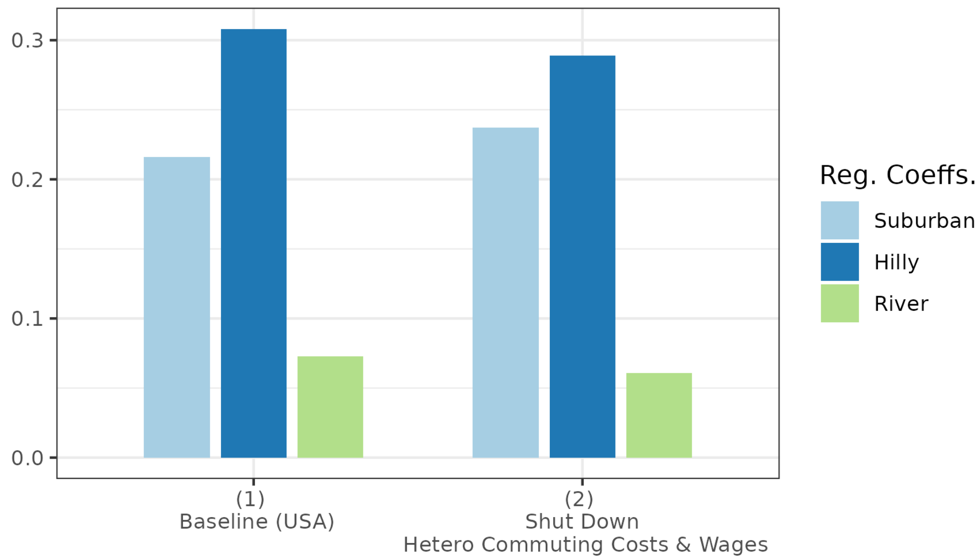
Note: This figure presents the estimated commuting semi-elasticities based on Equation (22). The left column shows the mean estimates and 95% confidence intervals separately for high- and low-income households—defined as above or below the city-specific median income—in developed cities (U.S. cities and Tokyo), and the right column shows those estimates in less-developed cities.

this counterfactual affects the income premiums observed in suburban, hilly, and river areas in U.S. cities.

Figure 11 presents the results of this counterfactual exercise. Following the format of Figure 9, we plot the estimated coefficients on the suburban, hilly, and river indicators from the regression specification in Equation (1). Column (1) shows the results using our extended model calibrated to U.S. cities. The baseline model replicates the empirical pattern that income tends to be higher in suburban, hilly, and river areas. When we eliminate heterogeneity in commuting costs and job access across income groups in Column (2), we find that these regression coefficients change only marginally. This result suggests that within-city variation in commuting costs and access to jobs across income groups plays a limited role in shaping the spatial distribution of residential income, and thus cannot account for the observed differences between developed and less-developed cities.

One may wonder whether our findings are consistent with those of Glaeser et al. (2008), who argue that within-city variation in commuting costs across income groups is a key driver

Figure 11: Counterfactual Simulation to Shut Down Heterogeneous Commuting Costs and Wages by Earning Potential



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 with heterogeneous commuting costs and wage rates. Column (2) reports the results where this heterogeneity is eliminated by setting $\overline{w}_{j,c}(L) = \overline{w}_{j,c}(H)$, *i.e., we assume that low- and high-earning-potential groups face the same commuting access.*

of the concentration of lower-income residents in central areas. They argue so by using a stylized monocentric city model, where all households are assumed to work at the city center. They also assume that the households differ only in terms of income, but not in terms of preference shocks, leading to an infinite elasticity of residential location choice. By contrast, our quantitative urban model incorporates two key departures from this setup. First, work locations are not fixed at the city center; instead, households choose among multiple employment locations, subject to commuting frictions. Second, households face idiosyncratic preference shocks over residential locations, which introduces a finite elasticity in location choice. These two features imply a more muted role for commuting-cost heterogeneity in shaping the spatial distribution of income—compared to overall income levels or commuting costs as highlighted in the previous section.

6.4. Distributional Impacts of Development

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

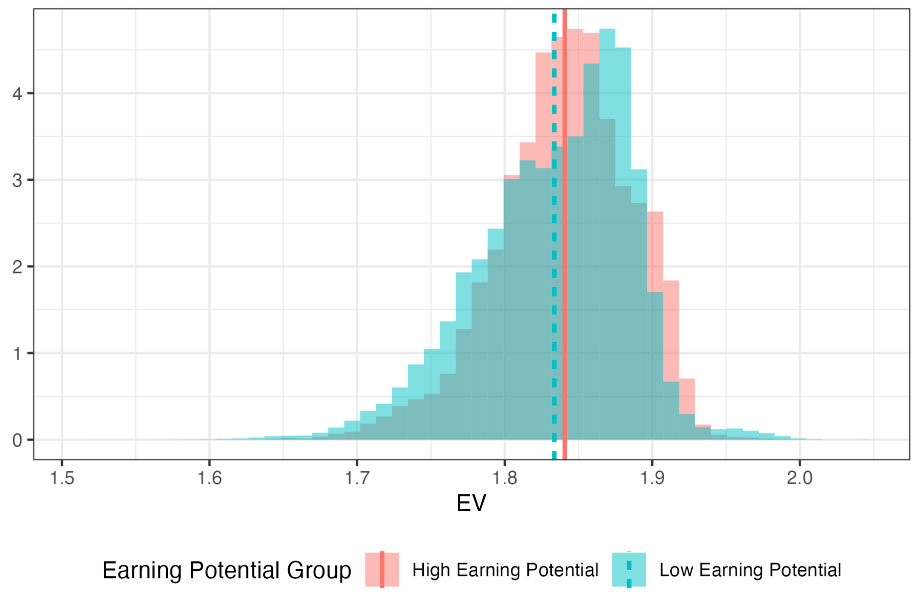
Specifically, using our model calibrated to U.S. cities, we compute welfare gains from a uniform increase in productivity – from the levels observed in less developed cities to the U.S. level, corresponding to a 2.0 log-point rise – as analyzed in Section 6.2. We examine how this aggregate productivity 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 homogeneous within earning potential types: the productivity increase uniformly raises both wages and housing costs across locations.

Figure 12 shows the distribution of equivalent variation (EV), in log points, across households by residential location and earning potential. The average EV is approximately 1.84 log points. 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.

We also find substantial dispersion in EVs across residential locations within each earning potential group, indicating significant spatial variation in housing cost changes, as we elaborate further below. Across earning potential groups, high-earning-potential households experience slightly larger average welfare gains than their low-earning-potential counterparts. This difference 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. The latter occurs because high-earning households increasingly sort into amenity-rich suburban areas as income rises, easing price pressure in urban cores. Our results suggest that the first force (i) modestly outweighs the second (ii), resulting in slightly lower welfare gains for lower-income households.

Figure 12: 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).

Table 8 reports how the EVs correlate with residential location characteristics, separately for high-earning-potential households (Columns 1 and 2) and low-earning-potential households (Columns 3 and 4). Columns 1 and 3 include only constant terms and therefore report average EVs, which are slightly higher for high-earning-potential households, consistent with the earlier discussion.

Across both groups, we find that EVs tend to be lower in suburban, hilly, and river-adjacent areas, and higher in neighborhoods with high job accessibility. This pattern aligns with the interpretation that, as overall city productivity rises, high-earning-potential households relocate from central neighborhoods to amenity-rich suburban areas. This reallocation reduces relative housing demand, and hence equilibrium housing costs, increases in the former areas, and decreases in the latter.

Table 8: 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		-0.043^{***}		-0.059^{***}
		(0.003)		(0.003)
Hilly		-0.044^{***}		-0.040^{*}
		(0.014)		(0.021)
River		-0.007^{***}		-0.012^{***}
		(0.002)		(0.002)
High Work Access		0.019^{***}		0.019^{***}
		(0.003)		(0.003)
Constant	1.841^{***}	1.868^{***}	1.834^{***}	1.846^{***}
	(0.002)	(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	0.411	0.000	0.453

Note: The 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.

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

We draw on new granular data from 127 cities in 26 countries to study how the spatial distribution of income within cities varies with economic development. We document that in less-developed countries, average incomes of urban residents decline monotonically in distance to the city center, whereas income-distance gradients are flat or increasing in developed economies. Neighborhoods with natural amenities – in particular hills and proximity to a river – are poorer than average in less-developed countries and richer than average in developed ones.

We develop a quantitative spatial model to evaluate potential explanations for these patterns, focusing on non-homothetic preferences over amenities, higher commuting costs or more centralized jobs in less-developed countries, and heterogeneity in commuting costs by household income level. We help discipline the importance of these channels by estimating commuting gravity equations in cities across the development spectrum.

Using the model calibrated to U.S. cities, we find that the differences in overall income levels and nonhomothetic preferences account for nearly half of the observed gaps in income premiums in suburban, hilly, and river neighborhoods between the U.S. cities and less-developed cities. 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. The remaining gap, which we leave as residuals, can be explained by lower amenity values and fewer public utilities in peripheral areas.

We also examine the distributional implications of overall city development. 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 households.

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," Tech. rep., Working Paper.

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

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015): "The Economics of Density: Evidence from the Berlin Wall," *Econometrica*.

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," Tech. rep., National Bureau of Economic Research.

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

Almagro, M. and T. Domínguez-Iino (2024): "Location sorting and endogenous amenities: Evidence from amsterdam," Tech. rep., National Bureau of Economic Research.

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, and B. Siddiqi (2020): "Transportation, Gentrification, and Urban Mobility: The Inequality Effects of Place-Based Policies," *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.

Bright, E., A. Rose, and M. Urban (2016): "LandScan Global 2015," .

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 (2023): "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.

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 (2022): "Structural change, land use and urban expansion," .

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.

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

Finlay, J. and T. Williams (2022): "Housing demand, inequality, and spatial sorting," *Inequality, and Spatial Sorting (April 28, 2022)*.

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*.

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," .

Harari, M. and M. Wong (2021): "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.

Hijmans, R. J. (2024): *terra: Spatial Data Analysis*, r package version 1.7-73, https://rspatial.github.io/terra/.

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*, Harvard Business School.

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," .

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.

LeRoy, S. F. and J. Sonstelie (1983): "Paradise lost and regained: Transportation innovation, income, and residential location," *Journal of urban economics*, 13, 67–89.

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*.

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," Tech. rep., 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.

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

Redding, S. J. and E. Rossi-Hansberg (2017): "Quantitative Spatial Economics," *Annual Review of Economics*, 9, 21–58.

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.

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): .

Tsivanidis, N. (2023): "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogota's TransMilenio," Unpublished Working Paper, U.C. Berkeley.

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: Cross-Country Evidence and Theory"

June 14, 2025

Peter Deffebach, David Lagakos, Yuhei Miyauchi, Eiji Yamada

Contents

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	420.1	21.8
Lima	2003	Peru	144490	388	2101.0	40.6
Managua	1998	Nicaragua	37082	92	257.6	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	873.5	25.7
Cairo	2001	Egypt	137513	429	2111.8	36.5
Damascus	1998	Syrian Arab Republic	38280	74	1941.1	22.8
Dar es Salaam	2007	U.R. of Tanzania: Mainland	26687	159	1176.7	29.0
Kinshasa	2018	D.R. of the Congo	42031	321	545.6	21.4
Mombasa	2015	Kenya	10868	32	190.2	10.5
Nairobi	2005	Kenya	20199	88	564.3	20.4
Nairobi	2013	Kenya	16794	102	594.9	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	1110.4	13.1
Cebu	2014	Philippines	28806	229	465.5	38.7
Chengdu	2000	China	31130	125	800.0	17.0
Colombo	2013	Sri Lanka	124673	376	1718.0	59.4
Da Nang	2008	Viet Nam	18171	50	401.1	14.3
Dhaka	2009	Bangladesh	70456	86	1090.2	22.4
Dhaka	2014	Bangladesh	118026	140	1844.5	29.7
Hanoi	2005	Viet Nam	63716	250	1131.0	31.9
Ho Chi Minh	2003	Viet Nam	102407	247	3967.2	40.5
Ho Chi Minh	2014	Viet Nam	42908	210	1877.5	35.5
Jakarta	2018	Indonesia	5000	534	1975.9	65.5
Karachi	2011	Pakistan	96507	170	1673.5	30.3
Kuala Lumpur	1999	Malaysia	80545	222	2562.2	46.9
Lahore	2010	Pakistan	89414	188	2583.9	43.1
Manila	1996	Philippines	231838	220	1671.1	37.4
Phnom Penh	2000	Cambodia	17398	70	351.3	14.3
Phnom Penh	2012	Cambodia	42074	85	536.0	17.5
Viang Chan	2007	Lao People's DR	27630	33	330.7	11.0
Yangon	2013	Myanmar	42224	620	782.4	29.7

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

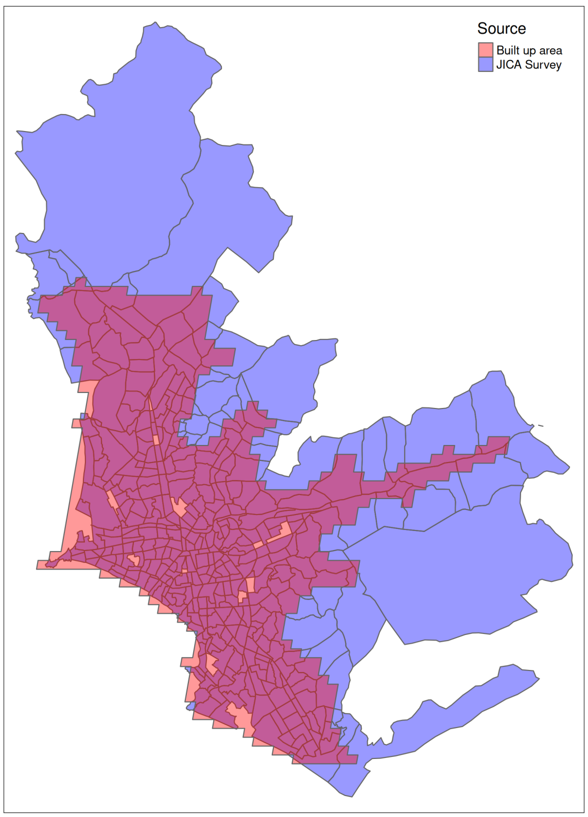
Country	Number of Cities	Total number of survey zones	List of Cities
Brazil	30	93876	Aracaju, Belo Horizonte, Belém, Campinas, Ceilândia, Cuiabá, Curitiba, Florianópolis, Fortaleza, Goiânia, João Pessoa, Londrina, Maceió, Manaus, Natal, Novo Hamburgo, Porto Alegre, Recife, Ribeirao Preto, Rio de Janeiro, Salvador, Santos, Sao Goncalo, Sao Jose dos Campos, Sorocaba, São Luís, São Paulo, Teresina, Uberlândia, Vila Velha
France	7	5419	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	27655	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.

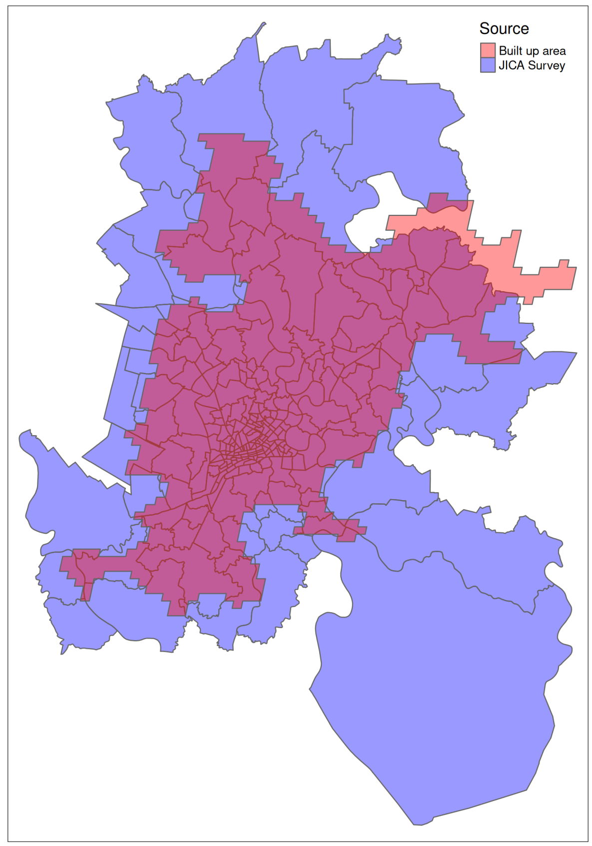
Figure A.1: Locations of Cities in our Data Set

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

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



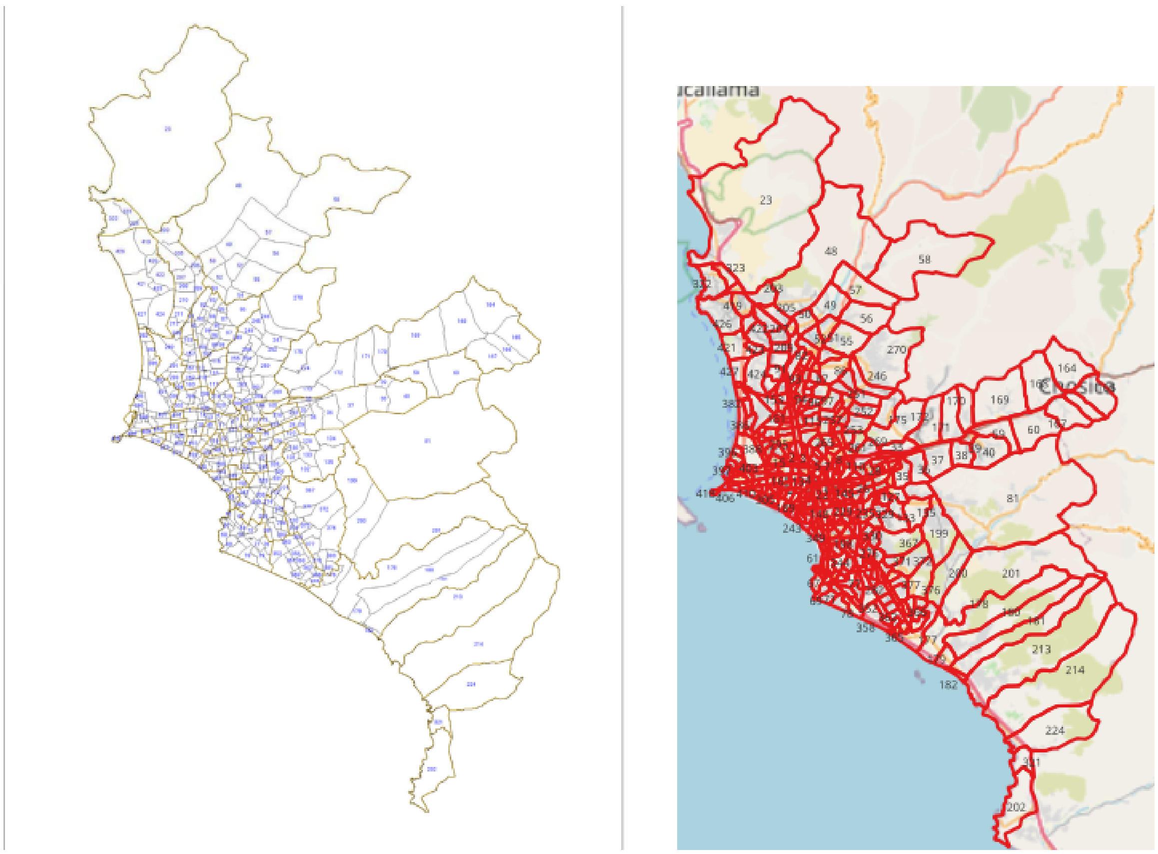
(a) Lima, Peru



(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: Left panel shows the screenshot of survey zones from a report for a travel survey. Right panel shows the screenshot of geocoded shapefiles of survey zones.

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

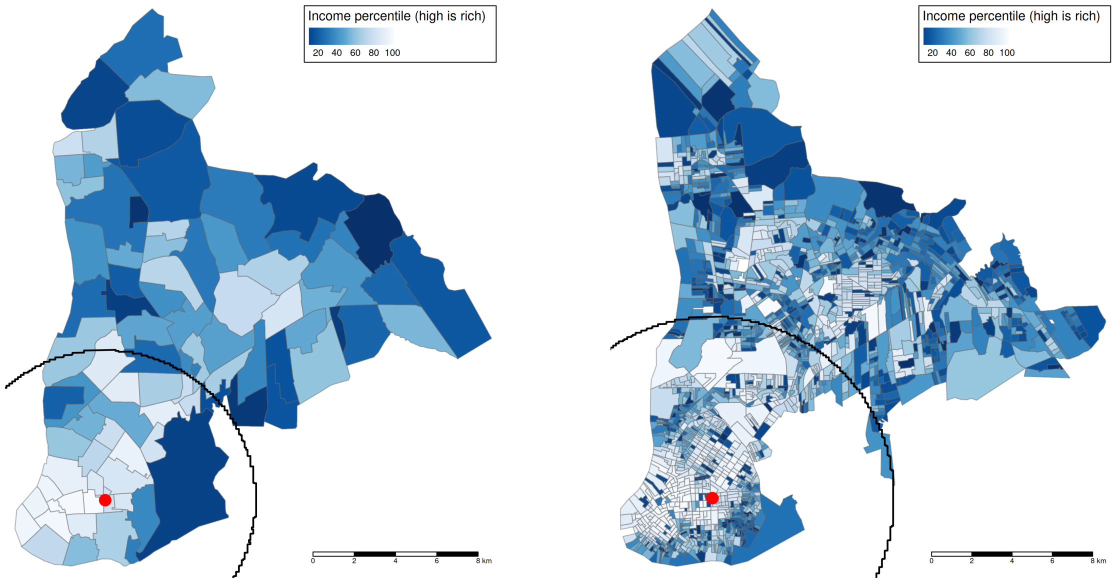
In this section, we provide a cross-validation of our residential income data from travel survey data and census data in Belem, Brazil, the only city in our data set where both types of data is available.

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 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 2km 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).

Finally, in Figure A.6, we directly compare the average neighborhood income between the JICA travel survey and the census data. Because the JICA travel survey is spatially coarser than those from the census, we aggregate the census income at the level of the travel survey. We compare the income percentiles between the two surveys at this level. We find remarkably tight relationships between these two data sets, with the majority of neighborhoods hewing to the 45-degree line, indicating identical income percentiles for both data sources.

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

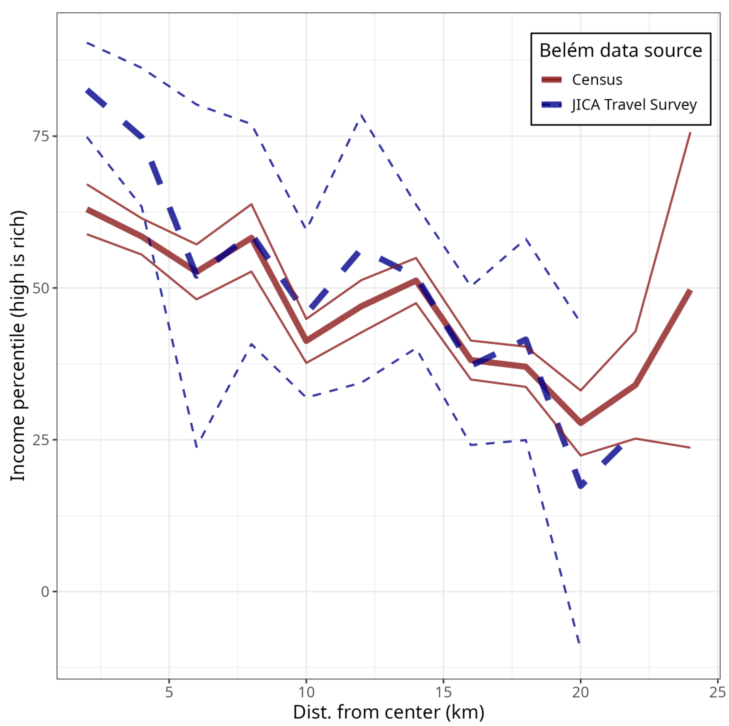
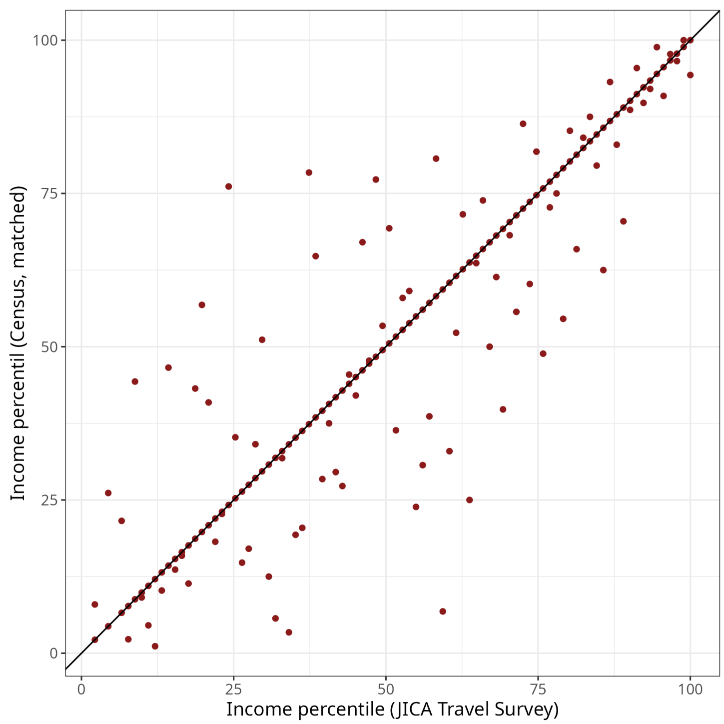


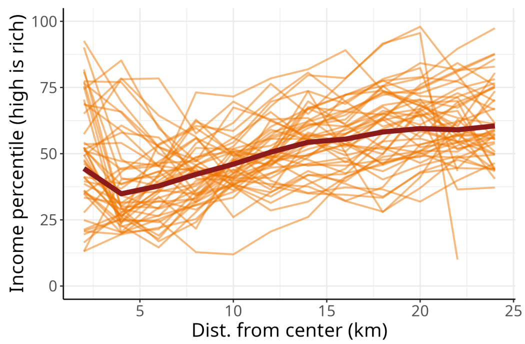
Figure A.6: Income ranking from census and JICA travel survey in Belem



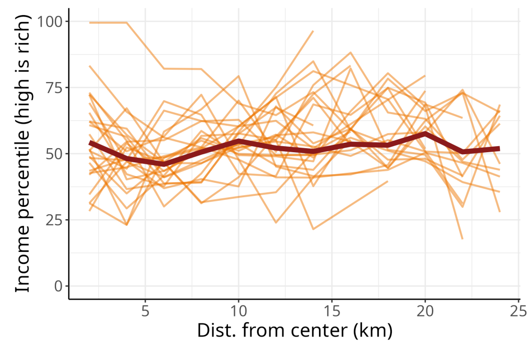
B. Additional Figures and Tables describing Income Patterns

B.1. Residential Income and Distance to City Center

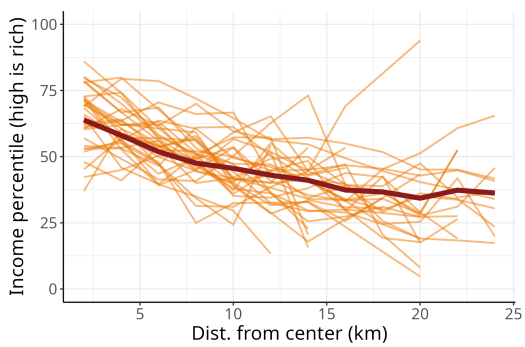
Figure B.1: Residential Income and Distance to City Center: Nonparametric Plots, by Continent



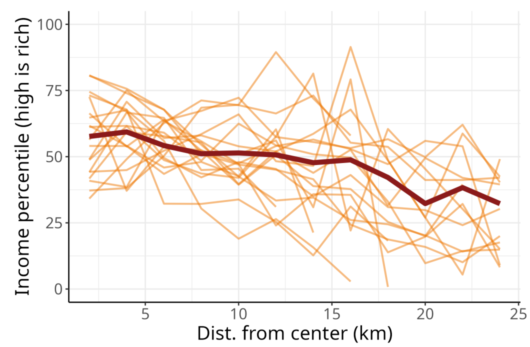
(a) USA



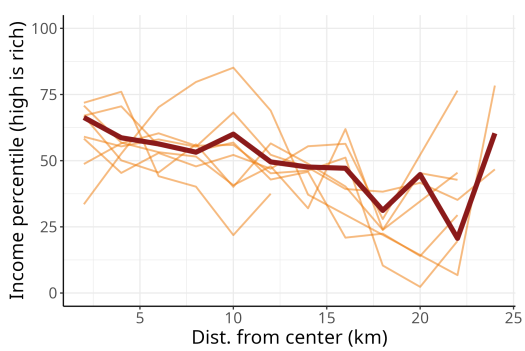
(b) W. Europe or Japan



(c) L. America



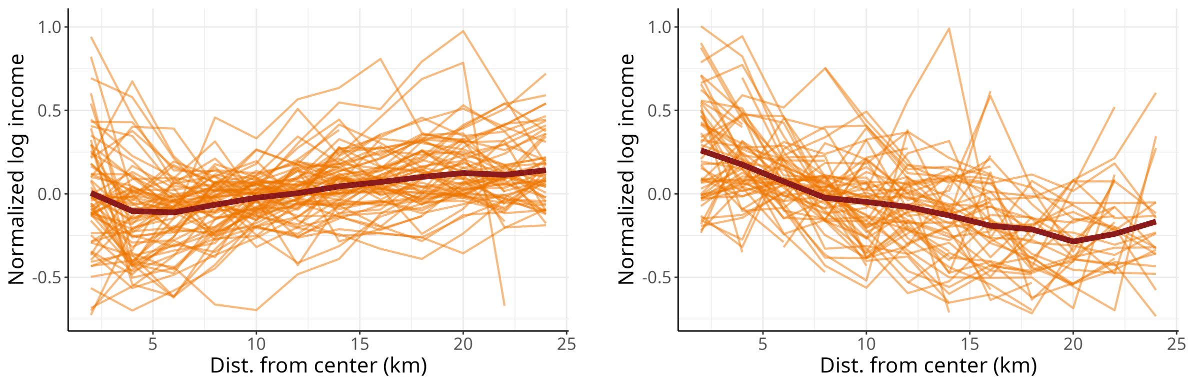
(d) Asia, E. Europe



(e) Africa, M. East

Note: Distance from the city center and average neighborhood residential income percentile. Emulates Figure B.1, grouping by continent rather than development status. Each light line represents a single city, and averages are highlighted in bold.

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

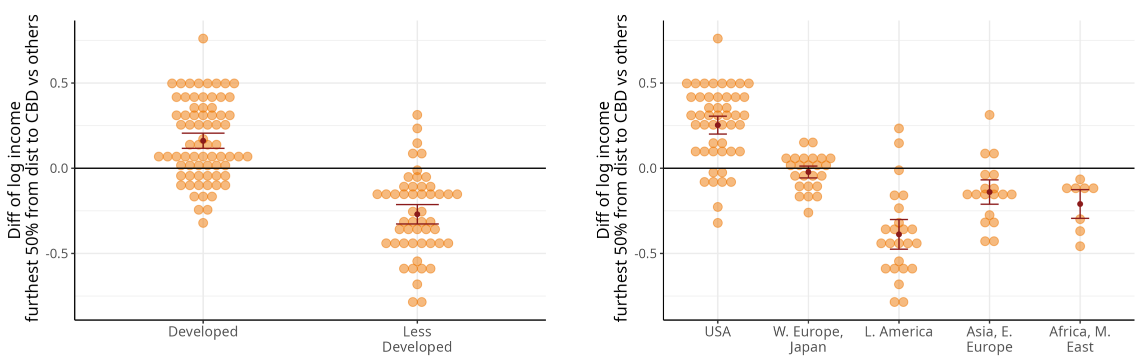


(a) Developed Cities

(b) Less Developed Cities

Note: Distance from the city center and income. Emulates Figure B.1, 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: Suburban-Urban Income Gap: log Income

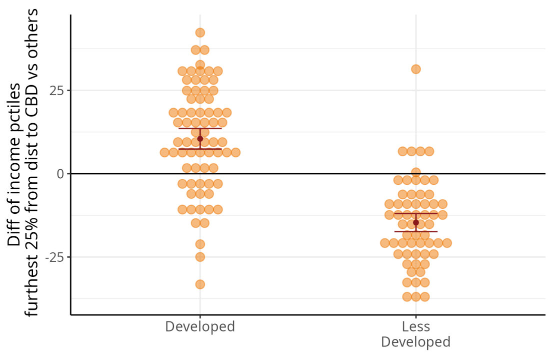


(a) By Development Status

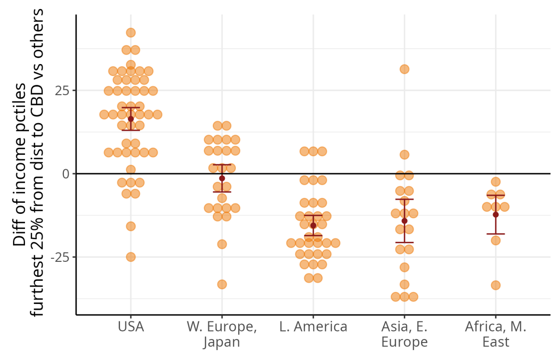
(b) By Continents

Note: Distance from the city center and income. Emulates Figure B.1, except using normalized log income on the y-axis, and with cities grouped by continent rather than development status. Each light line represents a single city, and averages are highlighted in bold

Figure B.4: Suburban-Urban Income Gap: By City, 25 percent as cut-off of ratio of income percentiles



(a) By Development Status



(b) By Continents

Note: The "suburban-urban income gap" of each city. Emulates Figure 4, except suburban neighborhoods are those with the outer 20 percent of the population, as opposed to the outer 50 percent as in our main specification. Each dot represents a city. Panel (a) groups cities by development status, Panel (b) by continent. We also plot the average value and its 95 percent confidence interval.

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

City	Country	Difference	Continent	Development Status
Hanoi	Viet Nam	-37.9	Asia, E. Europe	Less Developed
Ho Chi Minh City	Viet Nam	-37.1	Asia, E. Europe	Less Developed
Dhaka	Bangladesh	-35.9	Asia, E. Europe	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
Lima	Peru	-31.8	L. America	Less Developed
Curitiba	Brazil	-30.8	L. America	Less Developed
Goiânia	Brazil	-28.2	L. America	Less Developed
Phnom Penh	Cambodia	-28.1	Asia, E. Europe	Less Developed
Belém	Brazil	-28.0	L. America	Less Developed
Florianópolis	Brazil	-26.2	L. America	Less Developed
Porto Alegre	Brazil	-25.3	L. America	Less Developed
Seattle	United States	-25.0	USA	Developed
Vila Velha	Brazil	-24.8	L. America	Less Developed
Uberlândia	Brazil	-24.4	L. America	Less Developed
Ho Chi Minh City	Viet Nam	-23.9	Asia, E. Europe	Less Developed
Londrina	Brazil	-22.9	L. America	Less Developed
Campinas	Brazil	-22.0	L. America	Less Developed
Colombo	Sri Lanka	-21.6	Asia, E. Europe	Less Developed

(a) Bottom 20 Cities

City	Country	Difference	Continent	Development Status
Fresno	United States	42.3	USA	Developed
Las Vegas	United States	37.5	USA	Developed
Omaha	United States	36.7	USA	Developed
Buffalo	United States	32.7	USA	Developed
Detroit	United States	31.7	USA	Developed
San Antonio	United States	31.5	USA	Developed
Dà Nang	Viet Nam	31.4	Asia, E. Europe	Less Developed
Indianapolis	United States	30.5	USA	Developed
Tucson	United States	30.5	USA	Developed
Cleveland	United States	29.9	USA	Developed
Bakersfield	United States	29.2	USA	Developed
Oklahoma City	United States	29.1	USA	Developed
Providence	United States	28.6	USA	Developed
Columbus	United States	27.0	USA	Developed
Baltimore	United States	26.0	USA	Developed
Milwaukee	United States	25.9	USA	Developed
Albuquerque	United States	24.9	USA	Developed
Kansas City	United States	23.8	USA	Developed
Philadelphia	United States	23.7	USA	Developed
Houston	United States	23.7	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.

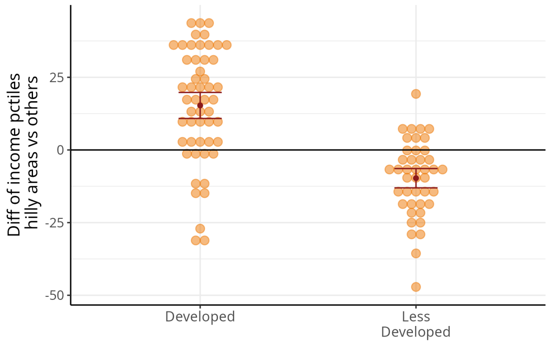
Table B.2: Suburban-Urban Income Gap: By Country

Country	Difference	Cities	continent	Development Status
Japan	-33.2	1	W. Europe, Japan	Developed
Peru	-31.8	1	L. America	Less Developed
Cambodia	-30.7	2	Asia, E. Europe	Less Developed
Sri Lanka	-21.6	1	Asia, E. Europe	Less Developed
Bangladesh	-20.6	2	Asia, E. Europe	Less Developed
D.R. of the Congo	-20.0	1	Africa, M. East	Less Developed
Viet Nam	-16.9	4	Asia, E. Europe	Less Developed
Myanmar	-15.8	1	Asia, E. Europe	Less Developed
Kenya	-15.2	3	Africa, M. East	Less Developed
Brazil	-15.1	31	L. America	Less Developed
Romania	-12.8	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
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
Philippines	-6.3	2	Asia, E. Europe	Less Developed
France	-5.4	7	W. Europe, Japan	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
United Kingdom	3.9	9	W. Europe, Japan	Developed
China	5.7	1	Asia, E. Europe	Less Developed
United States	16.4	48	USA	Developed

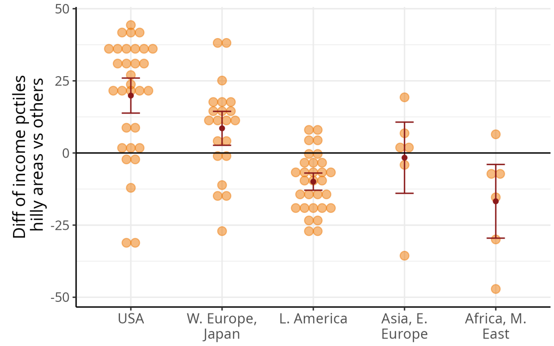
Note: A list of all countries in our neighborhood-level income dataset ranked by the gap ratio of income between suburban vs urban areas. Suburban neighborhoods are those containing the furthest 50 percent of the population.

B.2. Residential Income and Hills/Rivers

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



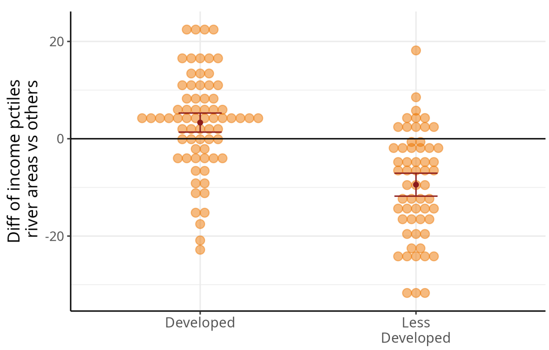
(a) By Development Status



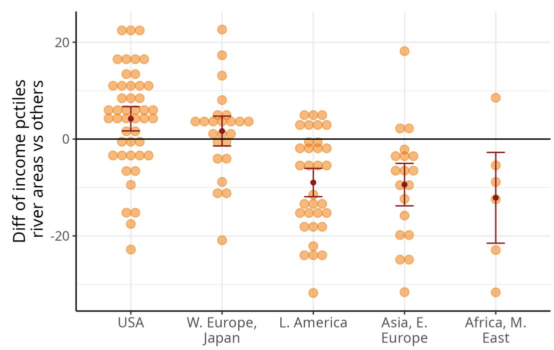
(b) By Continents

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) groups cities by development status, Panel (b) by continent. We also plot the average value and its 95 percent confidence interval.

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



(a) By Development Status



(b) By Continents

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) groups cities by development status, Panel (b) by continent. 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
Nairobi	Kenya	-47.1	Africa, M. East	Less Developed
Damascus	Syrian Arab Republic	-32.3	Africa, M. East	Less Developed
Dhaka	Bangladesh	-31.6	Asia, E. Europe	Less Developed
Lima	Peru	-29.2	L. America	Less Developed
Managua	Nicaragua	-25.2	L. America	Less Developed
Lahore	Pakistan	-25.1	Asia, E. Europe	Less Developed
Ho Chi Minh City	Viet Nam	-24.7	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
Belém	Brazil	-22.1	L. America	Less Developed
Colombo	Sri Lanka	-20.3	Asia, E. Europe	Less Developed
Vila Velha	Brazil	-20.1	L. America	Less Developed
Phnom Penh	Cambodia	-19.4	Asia, E. Europe	Less Developed
Newcastle upon Tyne	United Kingdom	-19.3	W. Europe, Japan	Developed
Maceió	Brazil	-18.5	L. America	Less Developed
João Pessoa	Brazil	-17.6	L. America	Less Developed
São Luís	Brazil	-17.3	L. America	Less Developed
Recife	Brazil	-16.9	L. America	Less Developed
Hanoi	Viet Nam	-15.8	Asia, E. Europe	Less Developed
Curitiba	Brazil	-15.8	L. America	Less Developed

(a) Bottom 20 Cities

City	Country	Difference	Continent	Development Status
Concord	United States	25.7	USA	Developed
Bakersfield	United States	23.3	USA	Developed
Honolulu	United States	22.6	USA	Developed
Lille	France	22.6	W. Europe, Japan	Developed
Las Vegas	United States	22.5	USA	Developed
Cleveland	United States	22.0	USA	Developed
Atlanta	United States	22.0	USA	Developed
Dà Nang	Viet Nam	21.9	Asia, E. Europe	Less Developed
Austin	United States	21.2	USA	Developed
Los Angeles	United States	18.3	USA	Developed
Sacramento	United States	17.7	USA	Developed
San Jose	United States	17.7	USA	Developed
Bradenton	United States	17.2	USA	Developed
Buffalo	United States	15.8	USA	Developed
Málaga	Spain	15.7	W. Europe, Japan	Developed
Seattle	United States	14.8	USA	Developed
Albuquerque	United States	14.2	USA	Developed
Indianapolis	United States	14.0	USA	Developed
Madrid	Spain	13.5	W. Europe, Japan	Developed
Lyon	France	13.3	W. Europe, Japan	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
Peru	-29.2	1	L. America	Less Developed
Nicaragua	-25.2	1	L. America	Less Developed
Kenya	-25.0	3	Africa, M. East	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
Bangladesh	-16.9	2	Asia, E. Europe	Less Developed
Pakistan	-14.6	2	Asia, E. Europe	Less Developed
Cambodia	-12.6	2	Asia, E. Europe	Less Developed
Myanmar	-12.3	1	Asia, E. Europe	Less Developed
Japan	-11.3	1	W. Europe, Japan	Developed
Côte d'Ivoire	-8.9	1	Africa, M. East	Less Developed
Brazil	-8.8	31	L. America	Less Developed
Malaysia	-7.5	1	Asia, E. Europe	Less Developed
Lao People's DR	-6.5	1	Asia, E. Europe	Less Developed
Viet Nam	-6.5	4	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
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
France	7.9	7	W. Europe, Japan	Developed
United States	8.0	48	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. Robustness of Regression Analysis

Table B.5: Further Robustness: Differences in Income Premiums in Suburban, Hilly, and River Neighborhoods between Less Developed versus Developed Cities

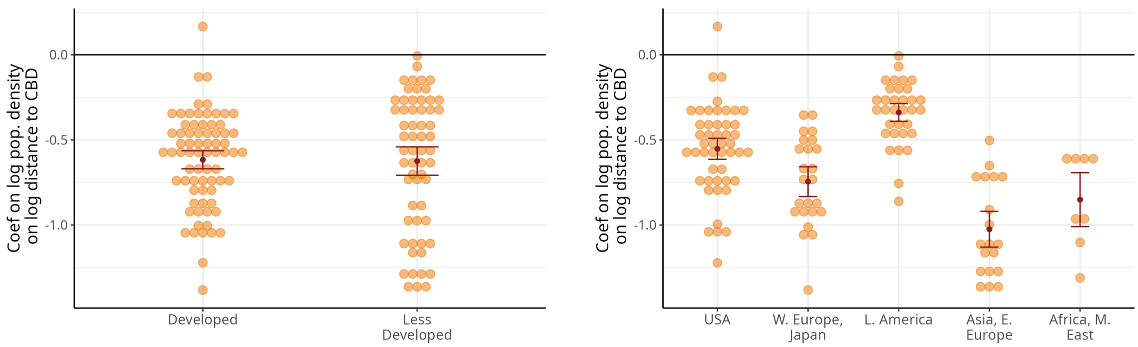
	Difference: Less Developed vs. Developed		
Specification	Suburban	Hilly	River
Main specification			
Income: Income percentile (high is rich)			
Distance: Top 50 percent distance from center			
Weight: Nbhd pop			
Income Measures			
Log income	-0.44 (0.04)***	-0.36 (0.06)***	-0.13 (0.03)***
Weighting schemes			
No nbhd pop weight	-24.2 (2.6)***	-21.1 (3.5)***	-9.6 (2.1)***
City subsets			
No Brazil	-22.9 (3.2)***	-17.0 (5.8)***	-10.4 (2.4)***
No USA	-12.2 (2.9)***	-15.4 (4.9)***	-7.3 (2.8)***
New World cities	-31.4 (2.8)***	-29.9 (3.4)***	-12.4 (2.4)***
Old World cities	-10.6 (3.7)***	-6.0 (6.2)	-7.5 (3.1)**
Exclude bottom/top 25% of cities by pop	-23.3 (3.8)***	-19.5 (5.7)***	-12.3 (4.0)***
Exclude bottom/top 25% of cities by area	-25.0 (3.5)***	-22.5 (4.3)***	-9.5 (2.9)***
Neighborhood subsets			
Exclude neighborhoods $\geq$ 15km of center	-18.7 (2.5)***	-19.0 (4.1)***	-8.7 (2.3)***
Exclude neighborhoods $\geq$ 20km of center	-22.6 (2.5)***	-20.3 (3.9)***	-9.0 (2.3)***
Distance measures			
Top 25 percent distance from center	-24.6 (2.4)***	-21.9 (3.6)***	-10.0 (2.1)***
Distance (km)	-1.8 (0.19)***	-21.6 (3.5)***	-8.4 (2.0)***
Log distance	-16.4 (1.8)***	-21.0 (3.5)***	-9.3 (2.0)***
Distance population rank (0-100)	-0.48 (0.05)***	-20.9 (3.6)***	-8.6 (2.0)***
Negative log population density	-11.6 (1.3)***	-14.1 (3.5)***	-5.1 (2.0)**
Negative log building volume density	-12.1 (0.95)***	-13.0 (3.7)***	-5.9 (2.0)***
City-level controls			
City population	-23.3 (2.7)***	-21.6 (3.5)***	-9.9 (2.4)***
City area	-23.5 (2.9)***	-18.8 (3.9)***	-10.4 (2.4)***
Ethnic diversity	-23.7 (2.6)***	-21.7 (3.4)***	-12.4 (2.4)***
Colonial group	-12.7 (3.5)***	-15.9 (4.0)***	-9.5 (2.8)***
Neighborhood-City controls			
Log area	-21.7 (2.4)***	-22.2 (3.5)***	-6.4 (1.8)***
Quadrant relative to center	-24.3 (2.5)***	-20.2 (3.6)***	-10.1 (2.1)***

Table B.6: Spatial Patterns of Population Density

Dependent Variable:	Log population density			
Model:	(1)	(2)	(3)	(4)
Variables				
$\text{Developed}_c \times \text{Suburban}_{j,c}$	-0.76***			-0.73***
	(0.05)			(0.05)
$\text{Less Developed}_c \times \text{Suburban}_{j,c}$	-0.82***			-0.73***
	(0.08)			(0.07)
$\text{Developed}_c \times \text{Hilly}_{j,c}$		-0.90***		-0.85***
		(0.11)		(0.10)
$\text{Less Developed}_c \times \text{Hilly}_{j,c}$		-0.42***		-0.37***
		(0.10)		(0.08)
$\text{Developed}_c \times \text{River}_{j,c}$			-0.70***	-0.65***
			(0.04)	(0.04)
$\text{Less Developed}_c \times \text{River}_{j,c}$			-0.98***	-0.82***
			(0.10)	(0.08)
Difference: Less $\textit{Developed}_c$ *vs* $\textit{Developed}_c$				
$\text{Suburban}_{j,c}$	-0.06			-0.003
	(0.09)			(0.08)
$\text{Hilly}_{j,c}$		0.47***		0.47***
		(0.15)		(0.13)
$\text{River}_{j,c}$			-0.28***	-0.18**
			(0.10)	(0.09)
Observations	145,393	145,377	145,393	145,377
City-Year FE	✓	✓	✓	✓
Unique City-Years	132	132	132	132
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.*

Figure B.7: Population Density and Distance to City Center

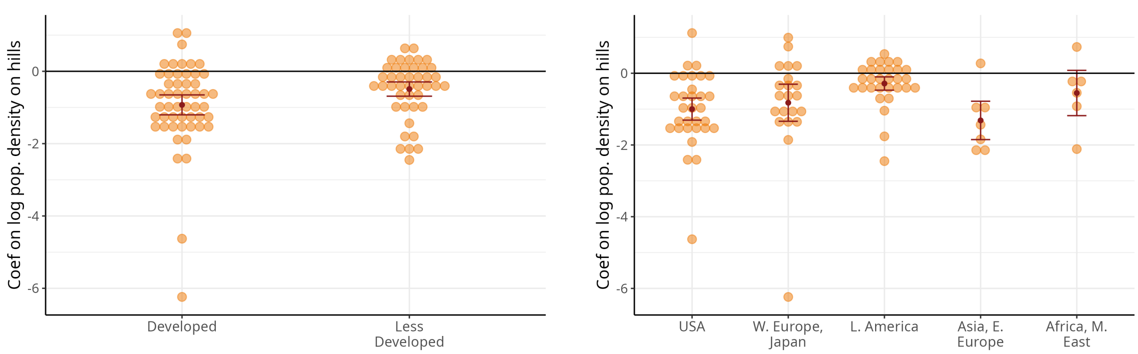


(a) By Development Status

(b) By Continents

Note: Population density gradient across cities. Emulates Figure 4, however instead of income ratios as the outcome, we analyze the analyze the population density gradient, defined as the coefficient of log distance to the city center regressed on log population density. Dots represent cities. Panel (a) groups cities by development status, Panel (b) by continent. We also plot the average value and its 95 percent confidence interval.

Figure B.8: Population Density and Hills

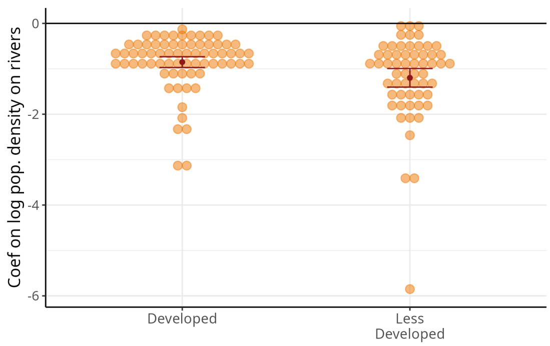


(a) By Development Status

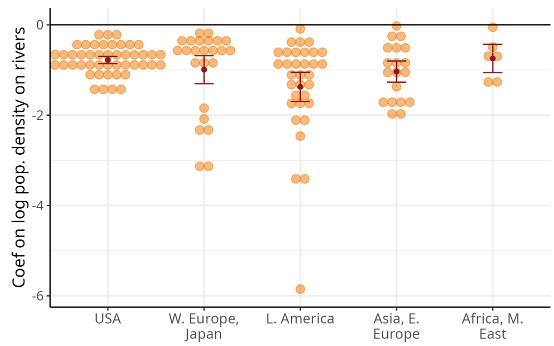
(b) By Continents

Note: Population density and hilliness across cities. Emulates Figure B.5, however instead of income ratios as the outcome, our outcome is the coefficient of hilliness when regressed on log population density. Dots represent cities. Panel (a) groups cities by development status, Panel (b) by continent. We also plot the average value and its 95 percent confidence interval.

Figure B.9: Population Density and Rivers



(a) By Development Status



(b) By Continents

Note: Population density and being near a river across cities. Emulates Figure B.6, however instead of income ratios as the outcome, our outcome is the coefficient of being near a river when regressed on log population density. Dots represent cities. Panel (a) groups cities by development status, Panel (b) by continent. We also plot the average value and its 95 percent confidence interval.

C. Additional Figures and Tables for Quantitative Analysis

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

City	Country	Commuting Semi-Elasticity (km)
Phnom Penh	Cambodia	0.52 (0.03)
Lahore	Pakistan	0.48 (0.02)
Chengdu	China	0.45 (0.02)
Vientiane	Lao People's DR	0.44 (0.02)
Hanoi	Viet Nam	0.33 (0.02)
Mombasa	Kenya	0.33 (0.02)
Dhaka	Bangladesh	0.30 (0.02)
Quezon City	Philippines	0.28 (0.01)
Abidjan	Côte d'Ivoire	0.23 (0.01)
Dà Nang	Viet Nam	0.22 (0.03)
Damascus	Syrian Arab Republic	0.22 (0.06)
Bucharest	Romania	0.22 (0.01)
Cairo	Egypt	0.20 (0.01)
Dar es Salaam	U.R. of Tanzania: Mainland	0.19 (0.01)
Managua	Nicaragua	0.18 (0.01)
Ho Chi Minh City	Viet Nam	0.18 (0.01)
Nairobi	Kenya	0.17 (0.02)
Lima	Peru	0.14 (0.00)
Cebu City	Philippines	0.13 (0.01)
Kinshasa	D.R. of the Congo	0.13 (0.01)
Kuala Lumpur	Malaysia	0.10 (0.00)
Yangon	Myanmar	0.10 (0.01)
Belém	Brazil	0.10 (0.01)
Karachi	Pakistan	0.09 (0.01)
Colombo	Sri Lanka	0.07 (0.00)
Jakarta	Indonesia	0.03 (0.00)

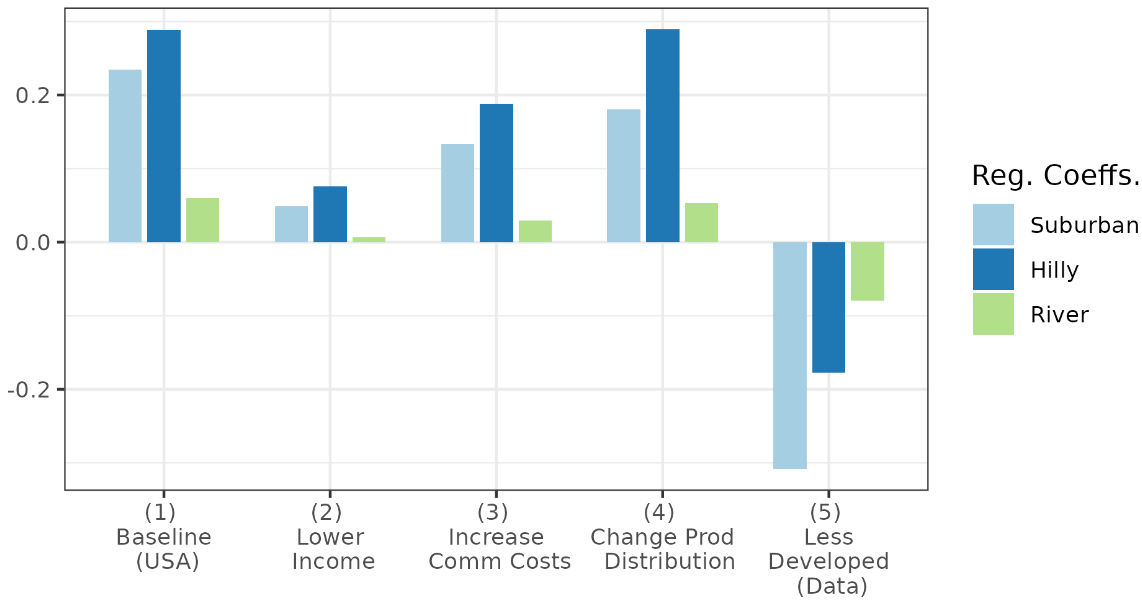
Note: Estimated semi-elasticity of road distance in kilometers using Equation (17). Parentheses indicate the standard errors, where the standard errors are clustered in two ways by origins and by destinations.

Table C.2: Regression Results of Estimated A_j and $\overline{w}_j$ by Income Groups

Dependent Variables:	$\log A^H_{j,c}(=\tilde{\psi}^H_{j,c}/\theta)$	$\log A^L_{j,c}(=\tilde{\psi}^L_{j,c}/\theta)$	$\log \overline{w}^H_{j,c}(=-\eta^H_{j,c}/\theta)$	$\log \overline{w}^L_{j,c}(=-\eta^L_{j,c}/\theta)$
Model:	(1)	(2)	(3)	(4)
Variables				
$\text{Developed}_c \times \text{Suburban}_{j,c}$	-0.12***	-0.08***	-0.08***	-0.07***
	(0.01)	(0.01)	(0.007)	(0.008)
$\text{Less Developed}_c \times \text{Suburban}_{j,c}$	-0.30***	-0.27***	-0.17***	-0.14***
	(0.02)	(0.02)	(0.02)	(0.02)
$\text{Developed}_c \times \text{Hilly}_{j,c}$	-0.26***	-0.26***	-0.03**	-0.04***
	(0.04)	(0.03)	(0.01)	(0.01)
$\text{Less Developed}_c \times \text{Hilly}_{j,c}$	-0.15***	-0.14***	-0.06**	-0.05
	(0.02)	(0.03)	(0.02)	(0.04)
$\text{Developed}_c \times \text{River}_{j,c}$	-0.02*	-0.07***	-0.02***	-0.02***
	(0.01)	(0.010)	(0.004)	(0.004)
$\text{Less Developed}_c \times \text{River}_{j,c}$	-0.17***	-0.16***	-0.08***	-0.06***
	(0.02)	(0.02)	(0.02)	(0.01)
Difference: Less Developed$_c$ vs Developed$_c$				
$\text{Suburban}_{j,c}$	-0.18***	-0.19***	-0.09***	-0.07***
	(0.03)	(0.02)	(0.02)	(0.02)
$\text{Hilly}_{j,c}$	0.11***	0.12***	-0.04	-0.01
	(0.04)	(0.04)	(0.03)	(0.04)
$\text{River}_{j,c}$	-0.15***	-0.08***	-0.06***	-0.03**
	(0.02)	(0.02)	(0.02)	(0.01)
Observations	25,259	25,367	25,463	25,463
City-Year FE	✓	✓	✓	✓
Unique City-Years	76	76	76	76
Weight neghborhoods equally within city	✓	✓	✓	✓
Subset	✓	✓	✓	✓

Note: The results of regression (1), where the outcome variables are estimated destination fixed effect from the commuting gravity equation (22), 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.

Figure C.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 6. 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* $\tilde{\kappa}_c$ *by 0.14, reflecting the average gap between U.S. and less-developed cities as shown in Figure 7. 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, as documented in Table 4. Finally, Column (5) displays the corresponding estimates for less-developed cities based on observed data, as reported in Table 2.*