Territorial Informality and Slums

Diogo Britto Lucas Conwell Breno Sampaio Gabriel Ulyssea
Lucas Warwar*

First Version: January 2025
This Version: March 27, 2026

Abstract

This paper investigates the defining characteristics of slums, the trade-offs their residents face, and the aggregate consequences of formalization policies. Using granular data from nearly 80,000 census blocks across Brazil's 19 largest metropolitan areas, we establish six facts showing that slums are fundamentally different from poor but formal areas: they are embedded in prime city locations, offering better access to formal jobs and schools at low rents, but at the cost of poorer-quality infrastructure and substantially higher crime. We develop a spatial general equilibrium model with territorial informality and informal housing markets, risk-averse workers, and endogenous amenity and productivity spillovers that rationalizes these patterns. Formalization policies targeting the downsides of territorial informality, such as crime, infrastructure, and amenities, raise aggregate output and welfare but exacerbate inequality. Enforcement of zoning regulations reduces the benefits of other formalization policies by limiting slums' ability to accommodate large numbers of low-income residents at low cost.

Keywords: Slums, informal housing, urban development

*Britto: Bocconi University, diogo.britto@unibocconi.it. Conwell: University College London, l.conwell@ucl.ac.uk. Sampaio: Universidade Federal de Pernambuco, brenorsampaio@gmail.com. Ulyssea: University College London, BREAD, CEPR, IZA and IFS, g.ulyssea@ucl.ac.uk. Warwar: Stanford University, warwar@stanford.edu. We thank Varsha Vejalla and Isabella Helter for outstanding research assistance. We are also thankful to Clement Imbert, Imran Rasul, Julieta Caunedo, Yuhei Miyauchi, and participants in seminars and workshops for useful comments. All errors are ours. Ulyssea gratefully acknowledges the funding received from the European Research Council (ERC) Starting Grant under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 852744).

1 Introduction

Slums constitute one of the distinctive features of developing countries' cities (e.g. Henderson et al., 2016, 2021; Michaels et al., 2021a; Gertler et al., 2025). By 2024, around 1.1 billion people or 25% of the world's urban population resided in slums, with 2 billion more expected in the next 30 years (UN, 2024). Despite the sheer magnitude of this phenomenon, we still have limited systematic, large-scale evidence on slums' defining characteristics and the livelihoods of their residents.[1] Therefore, fundamental questions remain unresolved: Are slums simply poor neighbourhoods with inadequate infrastructure and housing quality? Or do they possess distinctive characteristics that set them apart from other low-income but formal areas? Understanding the defining characteristics of slums and the trade-offs associated to them is crucial for evaluating the individual- and city-level consequences of these settlements and for designing effective policies.

This paper contributes to filling these gaps in two ways. First, we provide comprehensive evidence on slum characteristics and the central trade-offs their residents face by leveraging granular data sources covering nearly 80,000 census blocks from the 19 largest metropolitan areas in Brazil. The country is an important empirical setting, as it has the world's fourth largest urban population, with 16 million people living in slums (Harari, 2024). Our dataset covers nearly half of the urban population and span a 40-fold range in income per capita across census blocks. Second, we develop a spatial general equilibrium model that rationalizes the key empirical patterns we document. We use the estimated model to quantify the relative importance of different components of slums' characteristics in determining individuals' welfare in a spatial equilibrium accounting exercise. Moreover, we investigate the city-wide trade-offs associated with slums by simulating different prototypical slums policies and assessing their effects on total output, welfare, and inequality.

Slums are typically defined based on a range of characteristics, including tenure insecurity, inadequate housing, and poor infrastructure and services (e.g. UN-Habitat, 2003). In this paper, we emphasize *territorial informality* as slums' key defining trait: they are located on land with no formal and clear legal rights which are typically non-marketable, and thus particularly hard to formalize (e.g. Collier and Venables, 2014; Henderson et al., 2021). Empirically, we classify census blocks (*setor censitário*) as slums using the official designation developed by the Brazilian National Bureau of Statistics (IBGE) that requires a census block to have illegal land occupation, and at least one of the following: (i) irregular urban form (e.g., narrow alleys and irregular streets) or (ii) limited access to services and infrastructure.

Our analysis leverages a range of new administrative records and survey data to construct a granular dataset covering all 76,967 census blocks in Brazil's 19 largest metropolitan areas,

[1] See Ulyssea et al. (2025) for a review of the literature on slums. More broadly, there is limited evidence on residential patterns in developing country cities due to a lack of representative, granular data(Harari, 2024; Asher et al., 2026).

or around 80% of the national population. First, the Decennial Population Census provides individual-level demographics, measures of access to infrastructure and access to public goods. Second, we combine the national school census with a school-level educational quality index (IDEB) to construct distance-weighted measures of access to schools and school quality. Third, we geolocate hospitalization records from the Public Health System (SUS) to provide some of the first measures of crime victimization by residential neighbourhood in a developing country. Fourth, we use the RAIS matched employer-employee administrative data, which covers the universe of formal establishments and workers, to geolocate all formal jobs and compute a distance-weighted measure of commuter market access (CMA). Fifth, we merge in *CadUnico*, Brazil's welfare registry, which provides a near-census of the low socioeconomic status (SES) population with additional household-level information, such as composition, expenditures, and housing materials. Sixth, Google Open Buildings data supply finely gridded measures of building height and coverage.

We build on this new block-level database to establish six key facts about slums in Brazil. First, territorial informality does not necessarily imply a lack of state-provided infrastructure. Most slums have basic infrastructure—such as sewer and piped water—but suffer from poor quality: higher prevalence of open sewers, for example, and trash accumulating in the streets. Importantly, we document huge heterogeneity across both slums and formal neighbourhoods. Second, slum residents endure no worse housing conditions—measured using wall and floor materials—than low-SES households in formal areas. Instead, we show that nearly constant, limited height distinguishes slums from the formal built environment, consistent with, for example, Henderson et al. (2021) for the city of Nairobi.

Third, we use our new slum-level data to show that victimization rates are 87% higher among slum dwellers than among residents of poor formal areas within the same metropolitan area, with sharp discontinuities at the border. This is consistent with slums being areas of weak state authority and a lack of monopoly of force—as discussed in Lessing (2020)—which can lead to violent confrontations and a high incidence of crime (Monteiro and Rocha, 2017). Indeed, facts 1-3 show that while observable characteristics vary widely across slums, they share a fundamentally weak state presence, or "state void," which we argue ultimately stems from their territorial informality.

Fourth, we show that slums are pockets of state void embedded in the city proper, frequently surrounded by affluent areas with full state presence. In particular, we compute a block-level measure of "distance segregation" following Harari (2024) to show that, location-wise, slums resemble rich formal areas. As a result, we document sharp discontinuities at the border between slums and their neighboring formal blocks in average household income, formal job density, and infrastructure. Importantly, these discontinuities are not present in poor but formal blocks.

Consistent with their prime locations, Fact 5 shows that slums provide superior access to for-

mal jobs and high-quality public schools, relative to comparably well-off formal neighbourhoods. Sixth, we summarize the unique trade-off that slum residents face. Slums offer much higher access to formal jobs and schools than poor formal neighbourhoods while maintaining low rents. However, this access comes at the cost of higher victimization rates and worse infrastructure.

Based on these facts, we develop a spatial general equilibrium model with territorial informality that explicitly incorporates slums' key distinguishing features. Slums are characterized by higher victimization rates and different amenity-related factors such as access to schools and infrastructure. In both slums and formal areas developers can build floorspace at a convex cost. However, because of territorial informality there is a cost of expropriation in slums that is increasing in the amount of floorspace, thus imposing an effective limit on slums' vertical growth.[2] Nevertheless, due to the lack of zoning slum developers can build additional floorspace within a given height. Heterogeneous and, crucially, risk-averse individuals observe these characteristics and choose a residential location and, conditional o receiving a formal job offer, their workplace. If they do not receive a formal offer, they can work locally as self-employed. Firms are subject to agglomeration spillovers via their productivity, and residential amenities also have an endogenous component that depends on the "nearby" shares of high-skill residents.

We calibrate the model using our neighbourhood-level data and validate it against our key empirical facts. We draw on the literature and use observed neighbourhood characteristics to externally calibrate standard parameters, internally calibrate ten parameters central to slums' defining characteristics, and then invert the model to obtain neighbourhood-level amenities and productivities. Even though our six facts are not targeted in the calibration, the model is able to replicate them, including the sharp slum border effects observed in the data.

Finally, we use the estimated model to decompose slum residents' welfare into different components of the state void bundle and then assess the gains from prototypical formalization policies. The slum characteristics we emphasize explain around 60% of the slum penalty to utility. Slum residents suffer higher crime, poor infrastructure, and inferior amenities, but the lack of zoning restrictions boosts housing supply and prevents gentrification. We then compare the gains in output and welfare from reductions in slum crime, property rights (i.e., no developer expropriation), zoning, infrastructure, and amenity upgrades. Formalization virtually always increases aggregate output but exacerbates inequality. Policies that target the downsides of territorial informality, such as crime, infrastructure, and amenities, nonetheless benefit all residents, particularly when implemented without accompanying zoning reforms. Enforcement of zoning regulations and other measures that effectively constrain housing supply, however, threaten slums' unique advantage: their ability to accommodate large numbers of low-income residents at low cost.

[2]This cost is analogous to the "informality cost" faced by firms when deciding to operate in the formal or informal sectors (see Ulyssea, 2020, for a review of the literature).

This paper contributes to several strands of literature. First, we build on the literature that studies the determinants of slum formation (Henderson et al., 2021; Gonzalez-Navarro and Undurraga, 2023), redevelopment (Gechter and Tsivanidis, 2023; Harari and Wong, 2024), and public goods provision (Michaels et al., 2021b; Harari, 2024). In the context of Brazil, Cavalcanti et al. (2019) develop a macro model of housing choice and slum formation, while Ferreira et al. (2024) develop a model of structural change and slum formation to investigate whether slums are intergenerational stepping stones. Our quantitative spatial model (e.g. Ahlfeldt et al., 2015) highlights how micro-geographic variation in education quality, job access, crime, and land supply translate into workers' neighbourhood choices, agglomeration spillovers, and housing supply.

Second, we contribute to the literature on urban spatial structure in developing countries. Most recently, Deffebach et al. (2025) show that neighbourhood income rises with distance to the city centre in the US but falls in developing countries. We show that slums allow for an exception to this pattern, and our model rationalizes the existence of these low-income neighbourhoods embedded in prime locations.

Finally, this paper contributes methodologically to the quantitative spatial economics literature in the tradition of Ahlfeldt et al. (2015), Monte et al. (2018), and Heblich et al. (2020). Seminal developing-country applications include transport (Balboni et al. (2025); Tsivanidis (2023)) and firm informality (Zarate (2024)). We start with a full state-of-the-art quantitative spatial model of workers, firms, and housing developers and then, in order to rationalize the trade-offs imposed by territorial informality, incorporate risk aversion and crime, informal housing markets, and infrastructure-dependent amenities.

2 Background and Data

2.1 Brazilian Slums as "State Voids"

The literature typically defines slums as areas characterized by the following: (i) insecure land tenure or residential status; (ii) inadequate access to infrastructure (e.g. safe drinking water and sanitation); (iii) poor housing conditions; and (iv) under-provision of public goods (e.g. UN-Habitat, 2003; Cavalcanti et al., 2019; Alves, 2021).

In contrast, we emphasize *territorial informality* as the key characteristic that defines slums, which gives rise to a *state void*. Slums constitute an extreme form of land informality, as they lie on non-marketable land that cannot be easily formalized (e.g Henderson et al., 2021), which is often government-owned or an environmentally protected area. In these spaces, the state has, historically, given up the monopoly on the legitimate use of force and on the right to exercise its authority to enforce laws, individual rights, and formal rules as in the rest of the city (e.g.

De Souza et al., 2009; Lessing, 2020). While we focus our analysis on slums in Brazil, we argue that the state void applies to slums in Latin America more broadly, and individual components of the state void apply more broadly to slums in low- and middle-income countries.

Land informality naturally implies a lack of property rights over the land and built structures but may also rationalize other slums' characteristics. Weak (or "hollow") state presence leads to non-Weberian scenarios in which slums are dominated by slumlords or criminal organizations who control social and economic activities, characterizing criminal rather state governance (e.g. Arias and Barnes, 2017; Lessing, 2020; Uribe et al., 2025). This can generate violent confrontations between the police and criminal organizations—or between rival organizations—with extremely adverse consequences for residents (e.g. Dowdney, 2003; De Souza e Silva et al., 2008; Monteiro and Rocha, 2017; Arias and Barnes, 2017). Nevertheless, the organization of crime and its relationship with law enforcement vary substantially within and across cities (Lessing, 2020), potentially leading to substantial variation in observed levels of violence across slums.

Territorial informality also shapes slums' built environment and public good access. The relationship between lack of property rights and low-quality housing and investments has been extensively analyzed in the literature (e.g. Soto et al., 2000; Galiani and Schargrodsky, 2010; Henderson et al., 2021). Moreover, because governments do not officially recognize slums and rather see them as illegal or temporary, investments in the provision of basic infrastructure and public services can be quite limited (UN-Habitat, 2003; Gulyani and Talukdar, 2008). Nevertheless, a lack of state authority does not necessarily imply a lack of state presence. Infrastructure provision depends on a myriad of factors, including the political importance of a given area or existing complementarities between state presence and criminal governance (Blattman et al., 2025; Uribe et al., 2025). Conceptually, then, one should expect great variation in observed characteristics across slums and over time, with territorial informality as the common denominator.

2.2 Data

The basic geographic unit of our empirical analysis is the census block, which is a more granular unit than the census tract in the US (Belchior et al., 2025). We use the operational definition of slums adopted by the Brazilian Bureau of Statistics (IBGE): a census block is classified as a slum, or "subnormal agglomerate," if it is characterized by illegal land occupation—unauthorized occupation of private or public land—which necessarily implies that residents lack property rights over land and built structures. Additionally, slum blocks must display at least one of the following: (i) irregular urban form (narrow and irregular streets, uneven plots, and unauthorized constructions), or (ii) limited access to essential public services, such as electricity, piped water, and sanitation (IBGE, 2024). Hence, this empirical definition perfectly aligns with the concep-

tual one discussed in the previous section that uses territorial informality as the defining trait of slums.

We follow this definition to create a dummy variable for all blocks that constitute slums for the 2000 and 2010 Censuses. We focus on the largest 19 Metropolitan Areas (MAs) of the country, which cover around 80% of the national population as well as a 40-fold range in income per capita across blocks—equivalent to the gulf between Niger to Luxembourg. We combine a range of data sources to characterize all slum and formal census blocks in these 19 MAs as well as their residents.

Household Surveys and Decennial Population Census The Decennial Population Census data allow us to geo-locate individuals at the census block level. We can thus characterize census blocks in terms of the socio-demographic characteristics of their residents, housing characteristics, access to basic infrastructure and public services, such as access to piped water, bathrooms, sewer, and garbage collection. We complement these neighbourhood-level measures with information from the National Household Survey (PNAD), which allows us to contrast slum dwellers' labor market outcomes with those of formal neighborhood residents across all urban areas in Brazil. Specifically, we consider labor market participation, unemployment, hours worked, formal/informal employment, and wages.

Education We combine rich data on school infrastructure from the School Census with school-level data from the Basic Education Development Index (IDEB), a measure computed by the Ministry of Education using information on graduation rates and performance on standardized tests. We combine these data with schools' geocoordinates, as recoreded by the Institute of Applied Economic Research (IPEA), to create block-level measures of access to school quality in 2007. Specifically, for a given block i we take the inverse distance-weighted sum of the IDEB index across blocks j within the same metro area, with weights given by $1/\exp\left(k \cdot \text{distance}_{ij}\right)$, with $k = 0.5$.

Victimization We use individual-level data on hospitalizations from the public universal health care system (SUS) which, crucially, reports the residential zip codes of patients. We geocode these addresses to obtain block-level measures of assault hospitalizations as a proxy for residents' victimization by violent crime.

Formal Employment We use the RAIS matched employer-employee administrative dataset collected by the Ministry of Labour comprising the universe of formal establishments and their formal workers. We use a restricted access version that contains the exact addresses of all formal establishments and jobs to compute a block-level measure of commuter market access

(CMA). Specifically, for any block i we compute the inverse distance-weighted sum of job counts across blocks j within the same metro area, with weights given by $1/\exp(k \cdot \text{distance}_{ij})$. We set $k = 0.0374$, which is equal to the product of the workplace choice elasticity and disutility of travel time from Tsivanidis (2023), and use straight-line distance.

Welfare Registry The Single Registry of Social Programs (*CadUnico*) is an administrative registry that centralizes information on all actual and potential beneficiaries of federal social programs, constituting a near-census of low socioeconomic status (SES) households in Brazil spanning from 2011 to 2020. It contains household-level information such as composition and characteristics of the dwelling (e.g. wall materials), and individual-level information on all household members including age, gender, schooling, year of birth, race, addresses, and income profiles of more than 110 million individuals.

Google Open Buildings We make use of the newly-released Google Open Buildings 2.5D Temporal Dataset, which provides building height and presence at a 4m-gridded level across most of the developing world. We calculate the mean share of built-up land, or "coverage," as well as the mean building height, conditional on coverage, for each block in 2016.

Geolocating Slum Dwellers' Addresses

As a result of slums' territorial informality, many dwellings do not have formal addresses or zip codes. Moreover, in surveys, slum dwellers without a zip code frequently provide the zip codes of the closest bordering formal area. To overcome this measurement problem, we develop a methodology to construct a mapping between slums and the zip codes most commonly used by their residents. For this crosswalk, we rely on the CNEFE (*Cadastro Nacional de Endereços para Fins Estatísticos*), which contains roughly 100 million households with their complete addresses, ZIP codes, and geographic coordinates. The CNEFE covers around 83% of slums in Brazil, thus providing us with a mapping between slum census blocks and the zip codes most frequently erroneously reported by their residents. We combine the CNEFE data with two additional sources: (i) official ZIP codes assigned by the federal government program *Mapa da Periferia*, which aimed at providing ZIP Codes to low SES communities and (ii) the open-source database *CEP Aberto*, which expands ZIP code coverage beyond the CNEFE.

Descriptive Statistics

These datasets provide us with information on 76,967 blocks in the largest 19 MAs in Brazil, of which 7,054 are slums. Table 1 shows the basic descriptive statistics of our data using three groups: all census tracts (column 1), slums (column 2), and poor formal tracts (column 3), where the latter are all formal tracts below the median average household income per capita.

We see this as the most comparable group to slums, as it is more feasible for them to be in slum dwellers' choice set relative to higher income formal areas.

Table 1: Summary Statistics: Tract-Level Characteristics

	(1) All	(2) Slums	(3) Formal Poor
A. Demographics & Income			
Population	857	845	976
	(407)	(459)	(470)
Monthly income per capita (Census)	949.0	333.7	363.5
	(1057.0)	(142.6)	(106.3)
Years of schooling, HH head (Census)	7.02	4.73	5.04
	(2.73)	(0.96)	(1.32)
Share living in rented hh (Census)	0.166	0.073	0.113
	(0.117)	(0.079)	(0.092)
Share informal/self-employed (Census)	0.432	0.405	0.457
	(0.083)	(0.103)	(0.104)
B. Neighborhood Access			
Commuter market access (×1,000)	1253.5	1376.9	789.1
	(1074.7)	(995.9)	(774.8)
Public school access	34.06	40.80	26.70
	(21.83)	(20.16)	(22.39)
IDEB school quality index	3.991	3.970	3.711
	(0.537)	(0.501)	(0.504)
Distance segregation (normalized)	0.781	0.743	0.869
	(0.183)	(0.148)	(0.176)
D. Infrastructure & Built Environment			
Infrastructure index	0.828	0.789	0.708
	(0.208)	(0.188)	(0.244)
Building coverage (2016)	0.437	0.513	0.376
	(0.211)	(0.224)	(0.234)
Mean building height (2016)	6.29	5.32	4.31
	(4.96)	(1.92)	(1.67)
Census tracts	76,967	7,054	27,922

Notes: Each column reports the mean across census tracts (setores censitários) in Brazil's 19 largest metropolitan areas. Standard deviations in parentheses. Column (1): all tracts. Column (2): tracts designated as *aglomerados subnormais* (IBGE). Column (3): non-slum tracts with below-median income per capita (population-weighted). Infrastructure index averages shares with piped water, sewer, bathroom, and garbage collection (Census). Distance segregation follows Harari (2024). Education and victimization measures are inverse-distance weighted. Sources: Census (2000), CadÚnico, RAIS, School Census/IDEB, SUS, Google Open Buildings.

3 Facts

We now leverage the breadth of the data sources discussed in the previous section to characterize slum dwellers and their livelihoods in urban Brazil. We do so by establishing six key facts regarding slums and their residents. In the reminder of the paper, we refer to "poor formal blocks" as those who have average household income per capita below the median—very few slums have average household income per capita that is above the median—while "rich formal" are those blocks above the median.

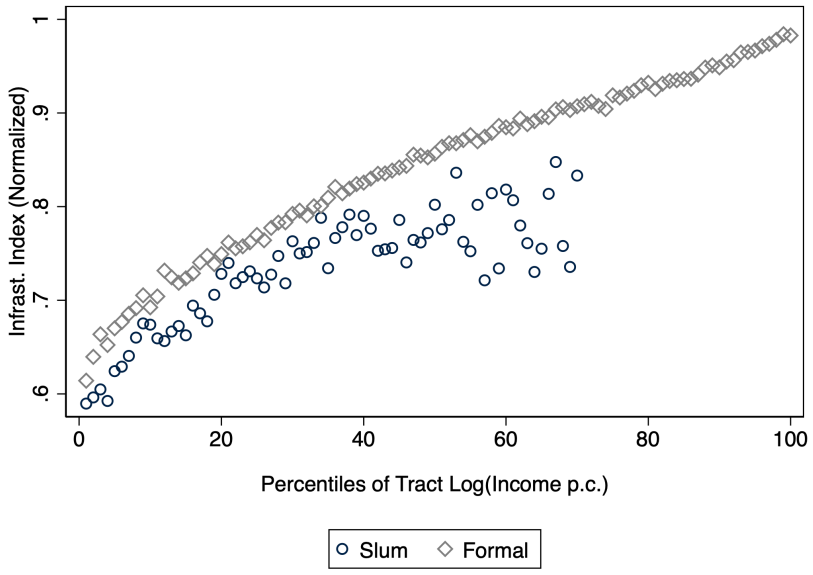
Fact 1: Territorial informality does not necessarily imply lack of state-provided infrastructure and services

We use Census data at the household level to compute a block-level infrastructure index that averages eight indicators: whether households have (i) access to piped water; (ii) access to a sewer system; (iii) access to trash collection service; (iv) street lighting; (v) paved streets; (vi) presence of sidewalks; (vii) trash accumulating in their street; (viii) open sewer in their street; and (ix) presence of trees. While items (i)-(iii) capture extensive-margin access to sanitation infrastructure and services, (iv)-(vi) measure the presence of basic urban infrastructure, and (vii)-(viii) capture the quality of the aforementioned services and infrastructure, and (ix) represents a standard amenity.

Figure 1 displays a binned scatterplot of this infrastructure index versus blocks' average income percentiles. At any level of income, slums have worse infrastructure relative to formal blocks, but the gap is quite small up to the 40th percentile and widens substantially above that. Figure 2 plots the average differences between slums and all formal areas (dark blue circle) and slums relative to poor formal areas (light blue circle) across the key components of the infrastructure index. The differences in the aggregate index turn out to mask important differences in individual infrastructural components. When compared to formal poor blocks, slums in fact have *higher access* to basic sanitation infrastructure—i.e. water and sewer—but *worse* quality as captured by higher presence of open sewers in the street, fewer paved streets and trees.

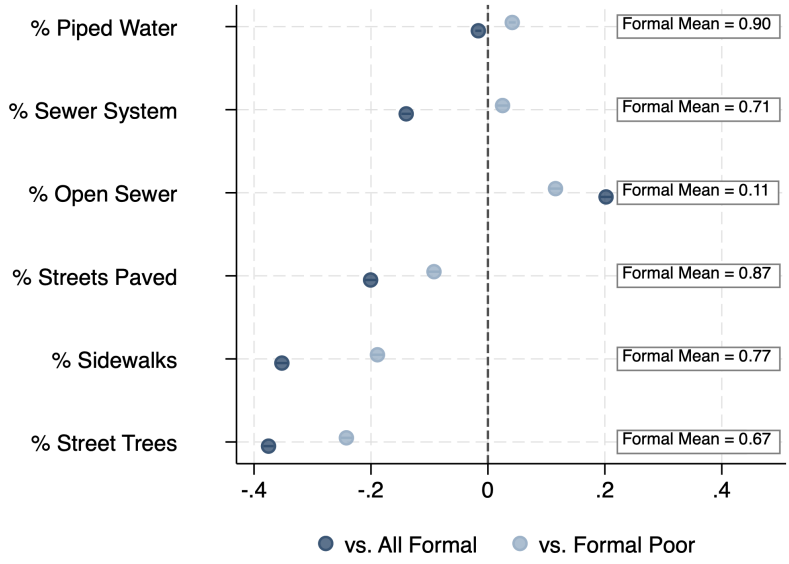
Finally, the distributions of the infrastructure index displayed in Figure 3 underscore the enormous heterogeneity in the infrastructure index across slums as well as poor formal blocks. Indeed, the two distributions almost perfectly overlap. Due to a combination of higher access but worse quality, slums do, on average, have slightly worse infrastructure than comparably-poor formal areas. However, slums cannot be adequately characterized simply by poor access to public infrastructure and services.

Figure 1: Block-level Infrastructure Index versus Neighbourhood Income



Notes: Figure displays a binned scatterplot with the average infrastructure index at the census block level on the vertical axis and percentiles of the country-wide distribution of block-level average income per capita on the horizontal axis, separately by slum status. Data from the Decennial Population Census.

Figure 2: Average Infrastructure Differences Between Slums and Formal Neighbourhoods

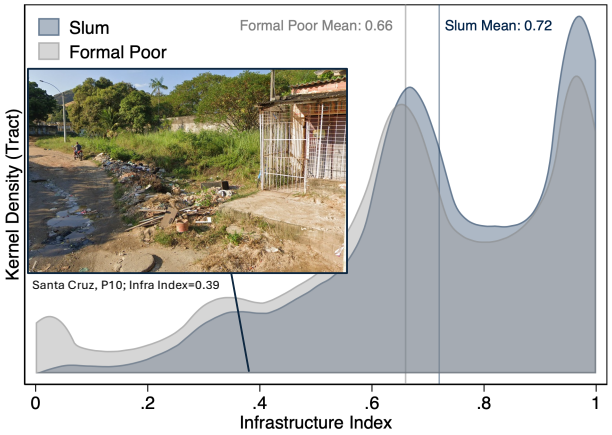


Notes: Figure displays the average difference between slums and all formal blocks (dark blue) and only formal poor blocks (light blue) across the main components of the infrastructure index. Data from the Decennial Population Census.

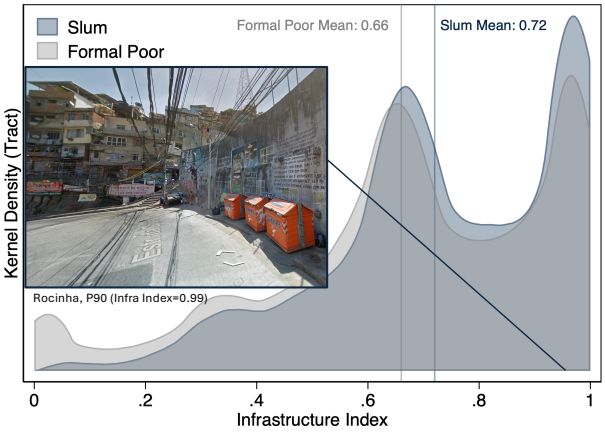
Fact 2: Slums are not characterized by worse housing conditions, but they do not grow vertically

In Table 2, we use information from *CadUnico*—which comprises a near-census of low-SES households—to assess the average differences in housing quality between slum- and poor formal-

Figure 3: Heterogeneity in Infrastructure Across Slums and Formal Neighbourhoods



(a) Example 1: Slum in the 10th percentile



(b) Example 2: Slum in the 99th percentile

Notes: Figures display the distribution of the infrastructure index across slums (dark blue) and formal poor blocks (light gray). Panels (a) and (b) repeat the same distribution but with pictorial examples of slums at the 10th and 99th percentiles of infrastructure quality, respectively. Data from the Decennial Population Census.

dwellers. We measure housing quality using two indicators of lower-quality materials: the share of households with cement or dirt floors and the share with uncoated brick walls.[3] Slums, on average, have a 5 percentage point *lower* share of households with cement or dirt floors and merely a 1.8 percentage point higher share of households with uncoated brick walls. Thus, both types of neighbourhoods have comparably low housing quality, and slum-dwellers do not suffer from worse housing conditions than comparable households in poor formal areas.

Figure 4, in turn, plots a binned scatterplot of mean building height versus percentiles of block-level average income, separately for slums and formal areas. Height clearly stands out as a distinctive characteristic of the slum built environment: slums' average building height remains at around 5 meters across the entire range of incomes of Brazilian census blocks. This finding dovetails strikingly with the evidence in Henderson et al. (2021) for the city of Nairobi, where 85% of constructions in slums are lower than 5 meters. By contrast, for formal areas in Brazil, building height increases in neighbourhood income, quite steeply so above the 70th income percentile.

Fact 3: Slums dwellers experience much higher violent crime victimization rates

As discussed in the previous section, slums are areas of weak state authority without rule of law. The duopoly of force and organized crime presence that characterize these state voids

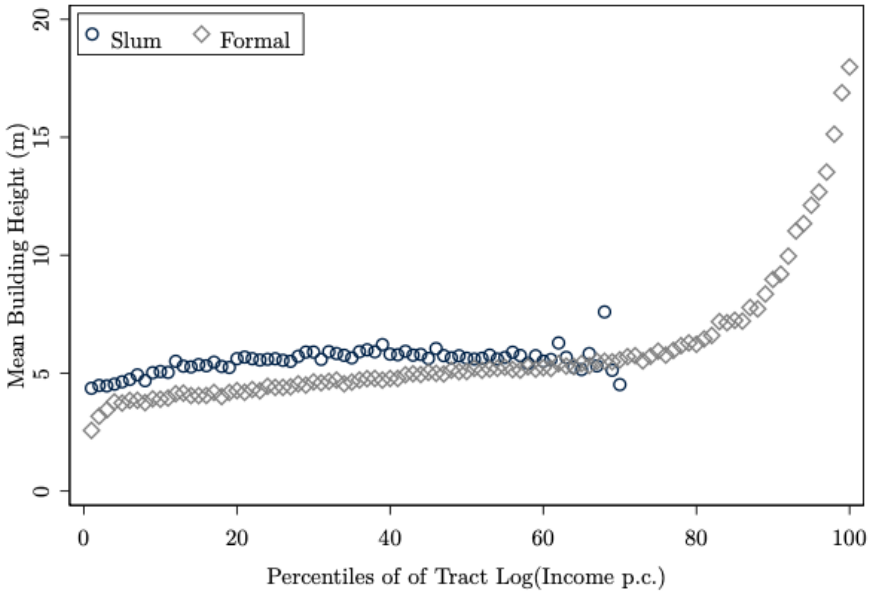
[3] *CadUnico* provides information on other types of housing materials, but they correspond to very small shares in the data.

Table 2: Slum Gap in Housing Conditions

	% Cement/Dirt Floor	% Uncoated Brick Wall
	(1)	(2)
Slum Dummy	-0.050***	0.018***
	(0.003)	(0.002)
Constant	0.459***	0.184***
	(0.002)	(0.001)
Observations	31,512	31,512

Notes: Robust standard errors in parentheses. *** indicates p<0.01, ** p<0.05, and * p<0.1. Table displays population-weighted regressions at the block level of indicators of low-quality housing materials, namely the share of households with cement or dirt floors and the share of households with uncoated brick walls, on a slum dummy and a constant. Data from *CadUnico*.

Figure 4: Building Height versus Neighbourhood Income



Notes: Figure displays a binned scatterplot of mean building height at the census block level on the vertical axis and percentiles of the country-wide distribution of block-level average income per capita on the horizontal axis, separately by slum status. Data on buildings' height from Google Open Buildings, and income from the Population Census.

could logically expose slum dwellers to higher violence and victimization rates. We examine this conjecture using the victimization measure described in Section 2.2, which aggregates hospitalizations due to assault from the National Health System. Table 3 shows that the average

slum has a 77% higher victimization rate when compared to all formal areas in Column 1; conditional on metropolitan area, in Column 2, this differential falls to 47%. However, in Column 3, compared to formal *poor* areas within the same MA in Column 3, the slum differential in victimization reaches a staggering 88%.

Table 3: Slum Gap in Victimization Rates

	Log(Victimization)		
	(1)	(2)	(3)
Slum Dummy	0.772***	0.466***	0.876***
	(0.036)	(0.023)	(0.029)
MA FE		✓	✓
Observations	67,399	67,399	28,216
R^2	0.005	0.395	0.442
Sample	All Blocks	All Blocks	Slums+ Formal Poor

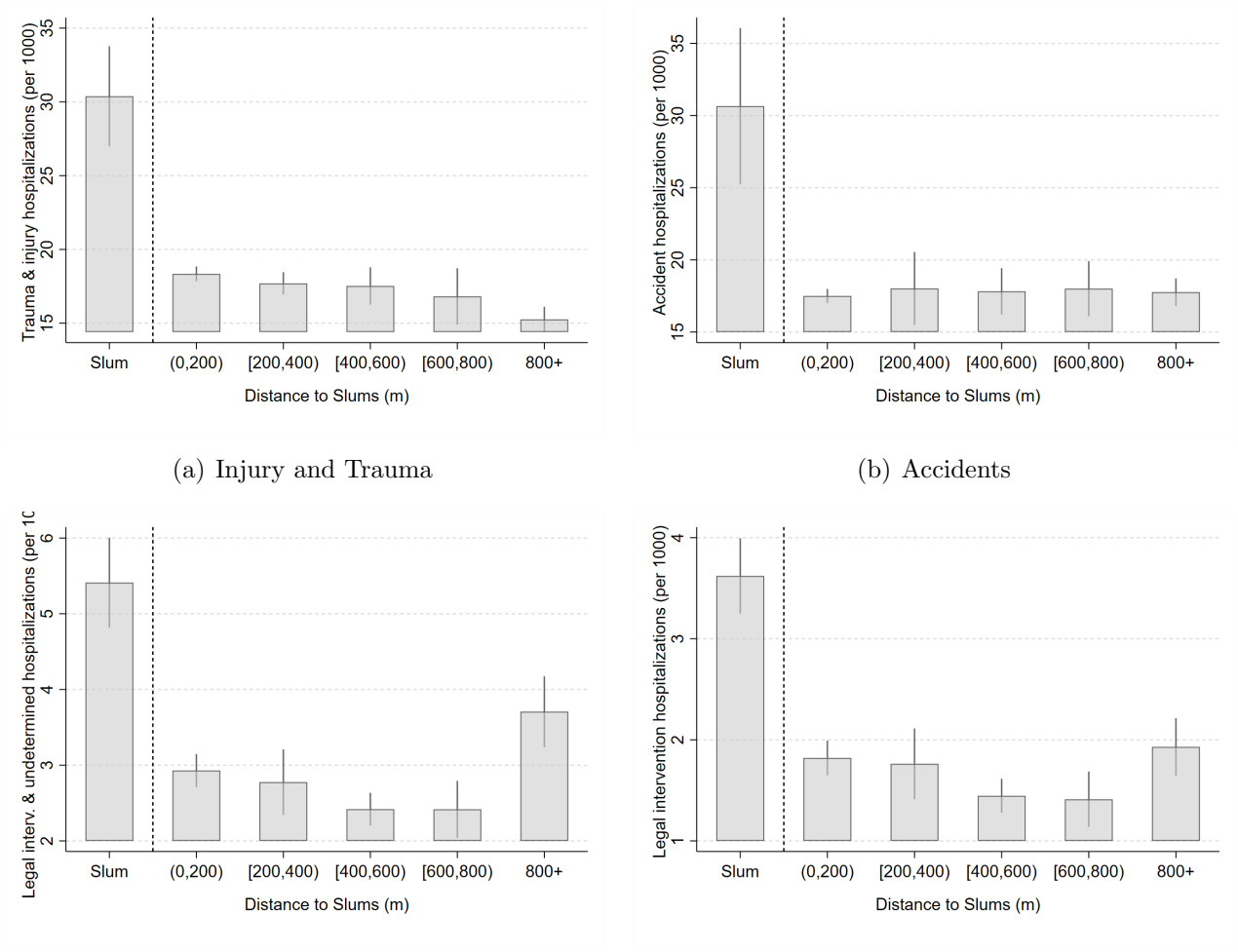
Notes: Robust standard errors in parentheses. *** indicates p<0.01, ** p<0.05, and * p<0.1. Table displays population-weighted regressions at the block level of log victimization rates, computed as described in Section 2.2, on a slum dummy and a constant. Data from SUS.

The differences in victimization rates grow even starker when we examine discontinuities at slums' borders. We examine assault hospitalization rates per 100,000 residents for four different components of our overall victimization measure: (a) hospitalization by injury and trauma; (b) accidents in transport, falls, etc.; (c) events of undetermined intent and, crucially, legal interventions by law enforcement agents or operations of war, including firearm injury by the police. Figure 5 plots the average victimization rates for slums and formal blocks, for the latter by five bins of distance to the nearest slum. Across all measures, slums have nearly double the victimization rates of immediately-adjacent formal census blocks.

Fact 4: Location-wise, slums resemble rich formal areas of the city, but with sharp discontinuities at the border

To assess each block's location within the city, we compute the *distance segregation measure* proposed by Harari (2024) at the census block level, which equals the average distance of a given block to neighborhoods in the top income quartile. Relative to the usual distance to the business district, distance segregation provides a more general measure of a location's attractiveness, particularly applicable to Brazil's relatively non-monocentric cities. Greater values of this measure correspond to more segregated blocks within the city. As Figure 6 shows, the distribution of the distance segregation measure across slums almost perfectly overlaps with that of formal

Figure 5: Discontinuities at Slums' Borders in Disaggregated Victimization Measures



(a) Injury and Trauma

(b) Accidents

(c) Undetermined Intent and Legal Interventions

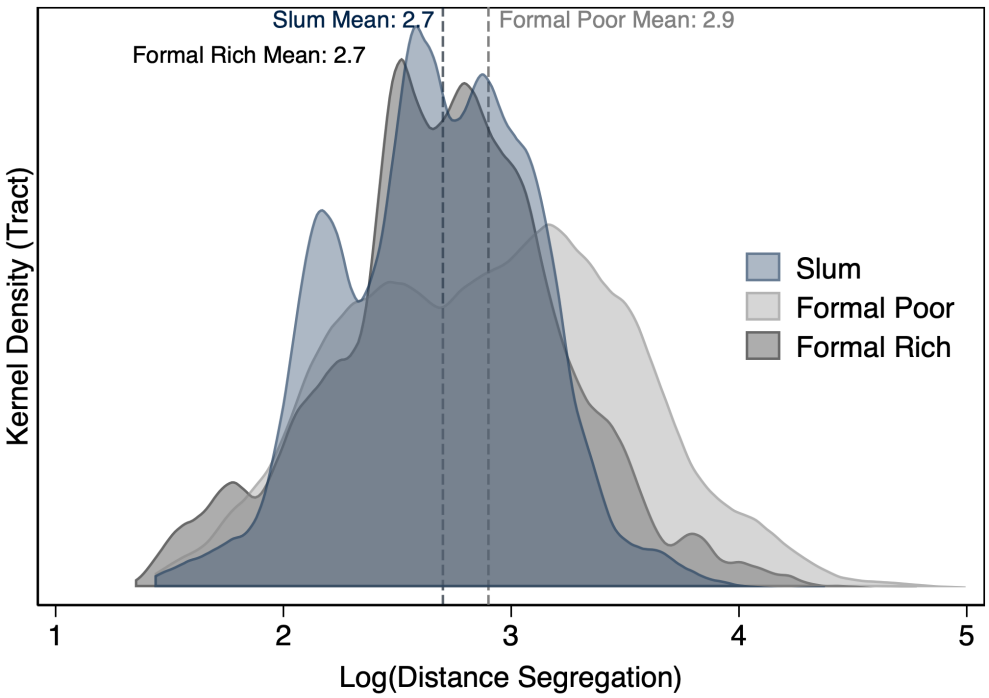
(d) Firearm Injury by Police

Notes: Figures display the mean assault hospitalization rates per 100,000 residents for four different components of our overall victimization measure for slums and for formal blocks, the latter disaggregated into five bins of distance to the nearest slum. Data from SUS.

rich neighborhoods. In contrast, formal poor neighborhoods are more distance-segregated—i.e. located in less-desirable locations—with a heavy right tail of highly-distance-segregated blocks.

Even though slums are geographically embedded and close to affluent areas in the city proper, their unique status as pockets of "state void" generates sharp discontinuities at their borders with formal neighborhoods. Figure 7 plots the average values of key neighborhood characteristics for slums and for formal blocks, by the same five bins of distance to the nearest slum. As seen in Panels (a)-(c), slums have substantially lower household income per capita, formal job density, and shares of households with sidewalks, when compared to immediately neighboring census blocks. Moreover, slum households are exposed to open sewers at three times the rate of residents of immediately-neighboring formal blocks.

Figure 6: Distance Segregation: Slums, Poor Formal, and Rich Formal

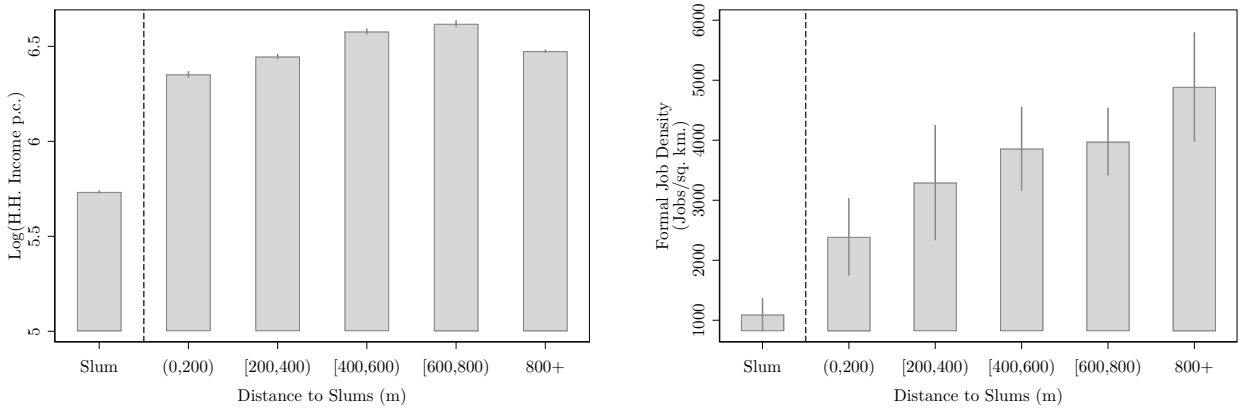


Notes: Figure displays the distributions of the Harari (2024) distance segregation measure, which equals the average distance from a given block to neighborhoods in the top income quartile, by neighborhood type: slum, poor (below-median-income) formal, and rich formal. Data from the Decennial Population Census.

Fact 5: Slums provide their residents with high access to formal jobs and good public schools

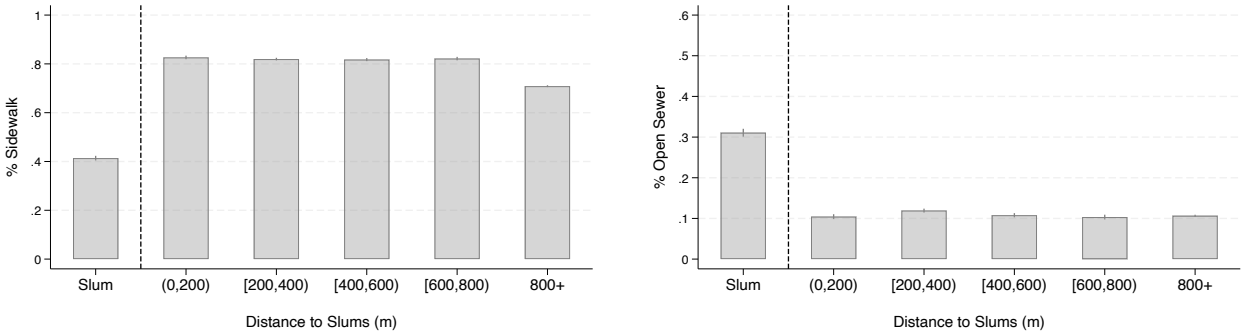
Consistent with Fact 4, we now establish that slums offer substantially greater access to formal jobs—i.e. greater commuter market access (CMA)—and public education. We measure both forms of access as inverse-distance-weighted sums of the counts of formal jobs and public schools across all census blocks. Additionally, we consider a measure of the quality of nearby public schools, namely an inverse-distance-weighted average of schools' national school quality index (IDEB). Panel (a) of Figure 8 shows that formal blocks mimic the typical pattern in developing-country cities, whereby CMA increases with neighborhood income (Deffebach et al., 2025). In contrast, slums display a flat relationship between neighborhood income and CMA, while displaying much higher CMA than formal blocks up until the 60th percentile of average household income per capita. In Panel (b), access to public schools varies little by neighborhood income, but, across the board, comparable-income slums benefit from substantially higher education access. Once we account for schools' *quality,* in Panel (c), slums' advantage over similar-income formal neighborhoods widens, once again, particularly in the bottom income percentiles.

Figure 7: Discontinuities at Slums' Borders in...



(a) ...Household Income Per Capita

(b) ...Formal Job Density



(c) ...% Households with Sidewalks

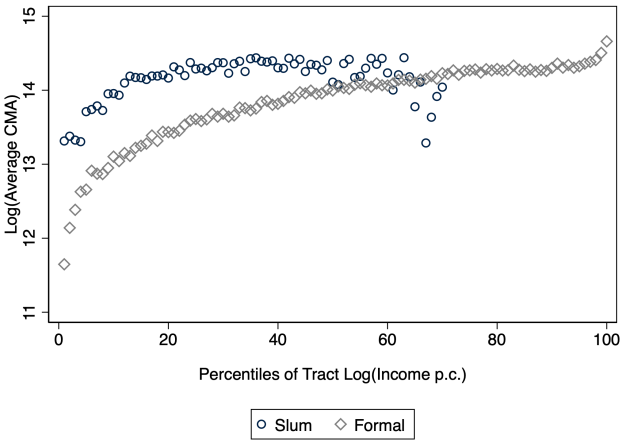
(d) ...% Households with Open Sewer

Notes: Figures display the mean household income per capita, formal job density, share of households with sidewalks, and share of households exposed to open sewers for slums and for formal blocks, the latter disaggregated into five bins of distance to the nearest slum. Data from the Decennial Population Census used in Panels (a), (c), (d); data from RAIS used in Panel (b).

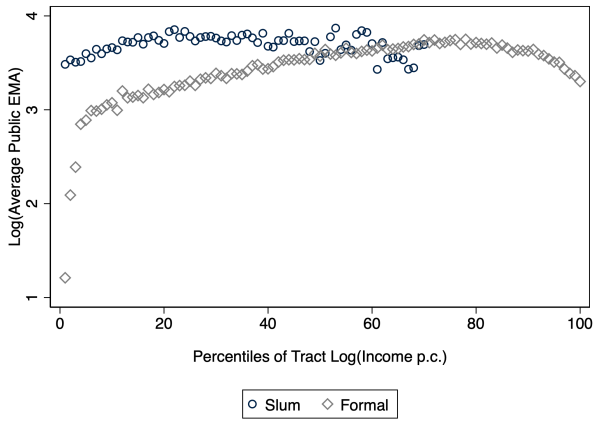
Fact 6: Slum dwellers face a unique tradeoff

Taken together, the preceding facts lay out the unique tradeoff which slum residents face. We summarize these upsides and downsides of slums in Figure 9, which plots the average difference between slums and poor formal areas across key dimensions. On average, slum dwellers benefit from 60% higher commuter market access (CMA) and 100% higher access to public schools — all at lower housing rents. However, these advantages come at the cost of exposure to slightly lower infrastructure and substantially higher victimization rates. As we have argued, this unique bundle of characteristics stems directly from slums' territorial informality, which fundamentally differentiate slums from poor but formal neighborhoods.

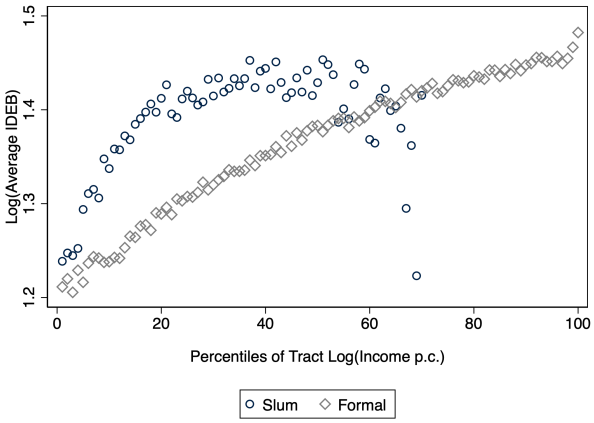
Figure 8: Neighbourhood Income and Access to...



(a) ...Formal Jobs (Commuter Market Access)



(b) ...Public Schools



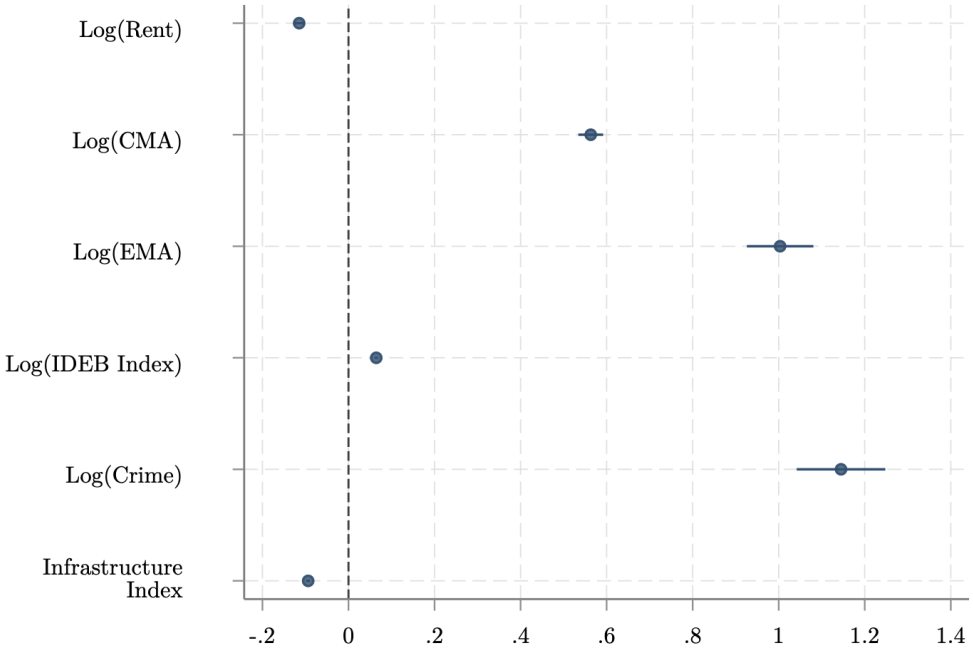
(c) ...Public Schools, Quality-Weighted

Notes: Figures display binned scatterplots with percentiles of the country-wide distribution of block-level average income per capita on the horizontal axis and, on the vertical axis, commuter market access to formal jobs (Panel a), inverse-distance-weightedaccess to public schools (Panel b), and an inverse-distance-weighted average of average of nearby schools' IDEB national school quality index (Panel c), in each case separately by slum status. Data from RAIS used in panel (a), and data from the School Census and IDEB Index used in panel (b).

4 A Spatial Model with Slums

We now develop a spatial general equilibrium model which incorporates the consequences of territorial informality. In particular, our model extends the canonical quantitative urban framework (Ahlfeldt et al., 2015; Tsivanidis, 2023) to incorporate key characteristics of slums: crime risk, informal housing markets, and public infrastructure.

Figure 9: Tradeoffs Faced by Slum Residents



Notes: Figure displays the population-weighted average difference between slums and formal poor (below-median-income) blocks in residualized housing rent, commuter market access (CMA), education market access (EMA, computed as the inverse-distance-weighted sum of block-level school counts), our inverse-distance-weighted school quality measure (IDEB), crime victimization rates, and our infrastructure index. We residualize housing rent, but not the other measures, on education, race, age, gender, state, and year dummies. Data on rents from the PNAD, data on CMA from RAIS, data on EMA from the School Census, data on the IDEB index from the corresponding database, data on crime victimization from SUS, and data on infrastructure from the Decennial Population Census.

Environment

We consider a static model of a closed city composed of N locations, or neighborhoods, classified into two neighborhood types: slum (S) or formal (F), henceforth indexed by g. Locations differ in their type-specific crime risk as well as the local probability of a formal job offer, infrastructure, school quality, and supply of developable land. In addition, locations offer heterogeneous amenity values and productivities, both of which comprise an exogenous and endogenous spillover component.

The model features three types of agents: workers, who choose where to live and, conditional on receiving a formal offer, where to work; firms, which produce a freely traded final good using labor and floorspace; and developers, who choose building height subject to convex construction costs. To capture the distinctive bundle of characteristics offered by slums, we depart from the canonical framework in three ways. First, workers are risk averse and face income risks arising from the probability of criminal victimization and job loss in the informal sector. Second, housing developers in slums face expropriation risk but can cram in additional floorspace. Third, the

exogenous component of residential amenities depends explicitly on neighborhood infrastructure and school quality. In equilibrium, rents and wages adjust to clear housing and labor markets in each location.

Workers

Fixed masses $\overline{h}_e$ of risk-averse workers of heterogeneous skill, indexed by e, populate our city; we denote the human capital level associated with a given skill level by χ_e. Workers first draw idiosyncratic preference shocks over home locations, indexed by i, and choose their place of residence with associated amenity B_i. They then draw a formal job offer with a home-location-specific probability λ_i and, if they receive an offer, draw idiosyncratic preferences over (formal) work locations, indexed by j, and choose their formal workplace. Iceberg commute costs τ_{ij} reduce formal earnings proportionately.[4] Workers without formal job offers work informally in their home neighborhoods for a fixed informal wage $\overline{w}^I$ and pay no commute costs.

All workers face income shocks that stem from the probability $\kappa_{g(i)}$ of being a crime victim, which varies by neighborhood type, i.e. slum vs. formal. Workers with informal-sector jobs additionally face an independent probability δ of job loss. Victimized workers or those who lose their jobs lose their labor income and instead receive a fixed government benefit $\underline{w}$. Workers have preferences over consumption of a homogeneous, freely-traded good, and housing. We assume preferences that nest non-homothetic Stone-Geary utility over housing and consumption inside a standard constant relative risk aversion (CRRA) utility function:

$$u_{ijse\omega}\left(c,h\right) \equiv B_i^{-\psi_e} \frac{\left[c^{1-\gamma}\left(h-\overline{h}\right)^{\gamma}\zeta_{j\omega}^{\mathbb{1}\{s=F\}}\right]^{1-\nu}}{1-\nu}\varepsilon_{i\omega}^{-1}. \tag{1}$$

where ω indexes workers, i denotes residential locations, s sectors, and j workplace location, c is the freely-traded final numeraire good, and h denotes housing.

To understand this specification of preferences, consider first the outer CRRA layer. The parameter ν in CRRA utility equals workers' relative risk aversion; in the typical case of $\nu \geq 1$, the resulting utility values are negative. We thus assume that amenities, in the first term in Equation (1) enter in reciprocal form, relative to the standard formulation. Workers of different skills vary in the degree ψ_e to which they value amenities. Next, consider the three components of the inner Stone-Geary layer. First, the parameter γ denotes the housing share, and the minimum housing requirement $\overline{h} > 0$ makes housing a subsistence good. Second, for the formally-employed, $\zeta_{j\omega}$ denotes the idiosyncratic work location preference shock. In the last

[4]We assume that commute costs are paid by workers in monetary terms but do not affect the human capital supplied to firms.

term, worker utility depends on an idiosyncratic home location preference shock $\varepsilon_{i\omega}$.[5]

We solve workers' sequential choices backwards, starting with the choices of goods and housing consumption.

Consumption Choice Workers choose consumption after their choices of home and work locations, and the realization of crime and potential job loss shocks. We first consider formal-sector workers of skill e who commute from i to j, and denote the wage per unit of human capital paid to workers of skill e by firms in location j by w_{je}. Formal workers face no job loss risk, and as long as they are not victimized by crime, they enjoy their full wage income net of commute costs, $\frac{w_{je}\chi_e}{\tau_{ij}}$. Let r_i denote the rent per unit of floorspace in location i. By the usual Stone-Geary results, individuals receive indirect utility, net of idiosyncratic preferences, given by

$$v_{ije}^F \equiv B_i^{-\psi_e}\left[\overline{\gamma}\left(\frac{w_{je}\chi_e}{\tau_{ij}} - r_i\overline{h}\right)/r_i^{\gamma}\right]^{1-\nu}/\left(1-\nu\right),$$

where $\overline{\gamma}$ denotes the usual Cobb-Douglas constant.

Formal workers who suffer a crime shock instead lose all labour income and must finance consumption and housing out of the neighborhood-invariant government benefit $\underline{w}$; we assume that the government additionally pays for the subsistence housing amount, such that the indirect utility of benefit recipients satisfies $v_{ie}^G \equiv B_i^{-\psi_e}\left(\overline{\gamma}\underline{w}/r_i^{\gamma}\right)^{1-\nu}/\left(1-\nu\right)$.

We now turn to informally-employed workers. Recall that informal workers who do not suffer a crime or job loss shock pay no commute costs out of their wage $\overline{w}^I$, so they receive indirect utility $v_{ie}^I \equiv B_i^{-\psi_e}\left[\overline{\gamma}\left(\overline{w}^I\chi_e - r_i\overline{h}\right)/r_i^{\gamma}\right]^{1-\nu}/\left(1-\nu\right)$. If informally-employed workers lose their jobs, suffer crime, or both, they receive the same government benefit $\underline{w}$ and associated $v_{ie\omega}^G$ as in the case of formal workers.[6]

Income Risk Crime victimization and informal job loss shocks to workers' income occur prior to consumption choice. Given the respective probabilities $\kappa_{g(i)}$ and δ, formal-sector expected utility, conditional on home location i and work j, and again net of idiosyncratic shocks, thus equals

$$EU_{ije}^F \equiv \overline{\gamma}^{1-\nu}\frac{B_i^{-\psi_e}}{r_i^{\gamma(1-\nu)}}\left[\left(1-\kappa_{g(i)}\right)\frac{\left(\frac{w_{je}\chi_e}{\tau_{ij}} - r_i\overline{h}\right)^{1-\nu}}{1-\nu} + \kappa_{g(i)}\frac{\underline{w}^{1-\nu}}{1-\nu}\right]. \tag{2}$$

[5]In order to ensure that higher shock draws correspond to greater location utility, these shocks, like amenities, enter in reciprocal form. Note also that we model both the home and work location shocks as *preferences,* so they do not affect the human capital supplied to firms, conditional on work location choice.

[6]Note that the risk of job loss and crime are independent, so the probability of suffering either or both shocks is $\kappa_{g(i)} + \delta - \kappa_{g(i)}\delta$.

Informal-sector workers additionally suffer the risk of job-loss. As a result, an informally-employed worker who lives in i has expected utility

$$EU_{ie}^{I} \equiv \overline{\gamma}^{1-\nu} \frac{B_i^{-\psi_e}}{r_i^{\gamma(1-\nu)}} \left[\left(1 - \kappa_{g(i)} - \delta + \kappa_{g(i)}\delta\right) \frac{\left(\overline{w}^{I}\chi_e - r_i\overline{h}\right)^{1-\nu}}{1-\nu} + \left(\kappa_{g(i)} + \delta - \kappa_{g(i)}\delta\right) \frac{\underline{w}^{1-\nu}}{1-\nu} \right]. \quad (3)$$

Crucially, the curvature introduced by CRRA utility—which is proportional to ν—will substantially depress expected utility in whichever neighborhood type due to crime risk $\kappa_{g(i)}$. However, this effect is magnified in slums because, as we show in Section 3, they have substantially higher crime rates. Hence, workers' location choices will respond strongly to the trade-off between this risk and any reward of living in slums.[7]

Workplace Choice After choosing a home location, workers receive a formal job offer with probability λ_i, which provides them with the opportunity to choose a formal workplace.[8] We assume that the idiosyncratic workplace preference shocks $\zeta_{j\omega}$ in Equation (1) follow a Frechet distribution with shape parameter θ. By the properties of the Frechet, $\zeta_{j\omega}^{1-\nu}$ follows a Weibull distribution with scale equal to unity and shape $\theta/(\nu-1)$, which implies that the expected value of formal employment in i is given by

$$\Omega_{ie}^{F} \equiv \Gamma\left(1 + \frac{\nu-1}{\theta}\right) \left[\sum_{j'} \left(EU_{ij'e}^{F}\right)^{\theta/(1-\nu)}\right]^{(1-\nu)/\theta}, \quad (4)$$

and probability that skill-e workers in i who receive a formal offer choose to work in j given by

$$\pi_{j|ie}^{F} \equiv \left(-EU_{ije}^{F}\right)^{\theta/(1-\nu)} / \sum_{j'} \left(-EU_{ij'e}^{F}\right)^{\theta/(1-\nu)}. \quad (5)$$

Workers who fail to receive the formal employment offer receive the expected utility of local informal employment, $\Omega_{ie}^{I} \equiv EU_{ie}^{I}$.

Home Location Choice In the model's outermost layer, workers first choose home locations based on the expected utility across formal and informal employment. We again assume that the corresponding idiosyncratic shocks, $\varepsilon_{i\omega}$, follow a Frechet distribution with shape parameter η.[9] This distributional assumption yields home neighbourhood choice probabilities, conditional

[7] Note that we have implicitly assumed that crime shocks to income occur after work and so, in the case of formally-employed workers, do not affect the human capital supplied to firms.

[8] We assume that all workers who receive formal offers accept them.

[9] Using the same results as for the workplace choice shock, $\varepsilon_{i\omega}^{-1}$ then follows a Weibull distribution with scale unity and shape η.

on workers' skill e, given by

$$\pi_{ie} = \frac{\left[-\lambda_i \Omega_{ie}^F - (1-\lambda_i)\,\Omega_{ie}^I\right]^{-\eta}}{\sum_{i'}\left[-\lambda_{i'} \Omega_{i'e}^F - (1-\lambda_{i'})\,\Omega_{i'e}^I\right]^{-\eta}}. \tag{6}$$

The expected location value in thus reflects a tradeoff between the "rewards" of a neighbourhood—such as high wages and amenities, or low rents—and the risks introduced by the crime, job loss, and formal job offer shocks. Finally, a worker of skill e enjoys ex-ante expected utility

$$\overline{U}_e \equiv \Gamma\left(\frac{\eta+1}{\eta}\right)\left\{\sum_{i'}\left[\lambda_{i'} \Omega_{i'e}^F + (1-\lambda_{i'})\,\Omega_{i'e}^I\right]^{-\eta}\right\}^{-1/\eta}. \tag{7}$$

Floorspace Supply

In order to capture direct implications of slums territorial informality, slums and formal neighborhoods have distinct building technologies and market structures. Recall that each neighborhood, whether slum or formal, offers a fixed supply of land denoted by $\overline{T}_i$, which is available to developers at a neighborhood-specific price. We assume that developers can build on a constant fraction $\overline{\pi}$ of land, often referred to as "coverage." Once built, floorspace can be used interchangeably either commercially by firms or residentially by workers.

Formal Neighbourhoods In formal neighbourhoods, we adopt a standard model of floorspace construction, as in Gechter and Tsivanidis (2023) or Harari and Wong (2024). Perfectly-competitive developers choose the height v_i^F up to which they build on the developable area $\overline{\pi}\overline{T}_i$ at convex cost $c\left(v_i^F\right) \equiv \Xi_0\, {v_i^F}^{\Xi_1}$. In consequence, given local floorspace rents r_i, developer profits per unit land satisfy $\Pi_i^F \equiv \left(r_i - \Xi_0\, {v_i^F}^{\Xi_1}\right)\overline{\pi} v_i^F$.

Slums Developers face the same cost function to build up, $c\left(v_i^S\right)$, but their technology and returns differ in two key ways. First, to capture the lack of zoning regulations in territorially-informal areas, we assume that each unit height built in slums yields an additional factor σ of floorspace—for example, because developers can build lower ceilings and thus additional floors. Second, to capture the risks inherent to lack of legal property rights due to territorial informality, we assume that developers in slums face of a probability $\mathrm{p}\left(v_i^S\right)$ of expropriation, which increases in the height built.[10] Expropriated developers lose their rental revenue but nonetheless pay the already-sunk building costs. Since we assume perfect competition and risk neutrality on the part of developers, their profits per unit of land in slums satisfy $\Pi_i^S \equiv$

[10] This cost is analogous to the "informality cost" that is typically used to model the costs faced by firms that decide to operate in the informal sector (see Ulyssea, 2020, for a review).

$\left\{\left[1-\mathrm{p}\left(v_i^S\right)\right]\sigma r_i - \Xi_0\, {v_i^S}^{\Xi_1}\right\}\overline{\pi} v_i^S$. Not only workers but also developers thus face a risk-reward tradeoff in slums: they can avoid zoning and thus earn higher yet more volatile returns.

We now derive the floorspace supply curves for each neighborhood type, given local rents r_i. Developers bid up land prices to ensure zero profits. In *formal neighborhoods,* we solve explicitly for developers' optimal height choices,

$$v_i^F = \left[\frac{r_i}{\Xi_0\left(1+\Xi_1\right)}\right]^{1/\Xi_1}, \tag{8}$$

which yield a supply of floorspace $H_i^F = v_i^F \overline{\pi}\overline{T}_i$.

In contrast, the optimal height v_i^S in *slums* satisfies the following first-order condition

$$v_i^S = \left\{\frac{r_i\sigma\left[1-\mathrm{p}\left(v_i^S\right) - v_i^S \mathrm{p}'\left(v_i^S\right)\right]}{\Xi_0\left(1+\Xi_1\right)}\right\}^{1/\Xi_1}, \tag{9}$$

which accounts for the expropriation risk $p\left(v_i^S\right)$ and yields a floorspace supply $H_i^S = v_i^S \sigma\overline{\pi}\overline{T}_i$ inflated by the additional floorspace factor σ. The extent to which slums offer more abundant floorspace and thereby lower rents thus depends on the relative strength of the zoning evasion versus expropriation effects, as captured by the floorspace factor σ and expropriation probability $\mathrm{p}\left(v_i^S\right)$, respectively. In equilibrium, rents determine the share ϕ_i of floorspace used for commercial purposes by firms.

Firms

The production side of our model hues closely to the standard quantitative urban framework, in that we model a *formal sector* with perfectly competitive output and input markets. We interpret the fixed *informal-sector* income $\overline{w}^I$ received by workers as self-employment and thus do not model informal production.

In each location j, a representative formal-sector firm produces a freely-traded final good using labour and floorspace according to a constant-returns-to-scale Cobb-Douglas production function with labour share μ and, recall, location-specific productivity Z_j. The labour input comprises a constant elasticity of substitution aggregator of human capital across skill types, with elasticity of substitution ξ. Thus, given inputs of commercial floorspace H_j^C and human capital X_{je} of each skill, the formal firm in location j produces output

$$Y_j^F \equiv Z_j \left[\left(\sum_e X_{je}^{\frac{\xi-1}{\xi}}\right)^{\frac{\xi}{\xi-1}}\right]^{\mu} \left(H_j^C\right)^{1-\mu}. \tag{10}$$

Given the local wage w_{je} per unit of human capital of skill e and rent r_j per unit of floorspace,

profit maximization implies a demand for human capital of skill e of

$$X_{je} = \frac{w_{je}^{-\xi}}{\sum_{e'} w_{je'}^{1-\xi}} \mu Y_j \tag{11}$$

and a commercial floorspace demand of

$$H_j^C = \frac{(1-\mu) Y_j}{r_j} \tag{12}$$

in location j.

Amenities and Productivities

Amenities in our model introduce a role for public infrastructure and, as common in the literature (e.g. Ahlfeldt et al., 2015), spillovers between neighborhoods. In particular, amenities B_i consist of an exogenous component $\overline{B}_i$, and an endogenous one given by the density of high-skill residents in surrounding areas, $\tilde{l}_{iH}$, raised to an amenity-high-skill elasticity Σ, or:

$$B_i \equiv \overline{B}_i \tilde{l}_{iH}^{\Sigma}. \tag{13}$$

The exogenous component, $\overline{B}_i \equiv \mathrm{g}_i(\iota_i, \varsigma_i)$, depends on neighborhood infrastructure ι_i and school quality ς_i as well as a neighbourhood-specific residual. The "nearby" high-skill density, in turn, equals an inverse-distance-weighted average of the density of high-skill population $L_{i'H}$ per unit land $\overline{T}_{i'}$. In other words, $\tilde{l}_{iH} \equiv \frac{\sum_{i'} \exp(-\rho_B \Delta_{ii'}) L_{i'H}/\overline{T}_{i'}}{\sum_{i'} \exp(-\rho_B \Delta_{ii'})}$, where $\Delta_{ii'}$ denotes the distance between neighborhoods i and i' and ρ_B controls the strength of the spatial decay of these amenity spillovers.

Similarly, productivity Z_j rises with an exogenous location fundamental $\overline{Z}_j$ and the (endogenous) nearby employment density, $\tilde{D}_j$, according to

$$Z_j \equiv \overline{Z}_j \tilde{D}_j^{\epsilon}. \tag{14}$$

Analogously to the case of amenities, agglomeration spillovers depend on an inverse-distance-weighted average of the density of total employment $E_{j'}$ per land area $\overline{T}_{j'}$, given by

$$\tilde{D}_j \equiv \sum_{j'} \exp(-\rho_Z \Delta_{jj'}) E_{j'}/\overline{T}_{j'},$$

where ρ_Z controls the strength of the spatial decay of agglomeration spillovers.

Aggregation and Equilibrium

Aggregates The choice probabilities in Equations (5) and (6) directly yield the total population by home and work location, skill group, and sector. The total population of skill e living in i and working in the formal sector in j is given by $L^F_{ije} \equiv \overline{L}_e \pi_{ie} \lambda_i \pi^F_{j|ie}$, where $\overline{L}_e$ denotes city-wide populations, and λ_i the location-specific formal job offer probability. The total population of skill e living in i and working in the informal sector, in turn, equals $L^I_{ie} \equiv \overline{L}_e \pi_{ie} (1 - \lambda_i)$. The total population of skill e living in i, which enters the amenity spillovers, thus satisfies $L_{ie} \equiv \sum_j L^F_{ije} + L^I_{ie}$. Finally, total (formal) employment in j, which enters agglomeration spillovers to productivity, is $E_j \equiv \sum_{i,e} L^F_{ije}$.

Market Clearing Labour markets clear when the formal-sector demand for human capital of skill e in location j equals the supply, or

$$X_{je} = \sum_i L^F_{ije} \chi_e. \tag{15}$$

Given the Stone-Geary assumption on housing demand and the aggregate location populations derived above, housing market clearing requires

$$\begin{aligned} \sum_{j,e} L^F_{ije} \gamma \left[\left(1 - \kappa_{g(i)}\right) \left(\frac{w_{je} \chi_e}{\tau_{ij}} - r_i \overline{h} \right) + \kappa_{g(i)} \underline{w} \right] + \\ \sum_e L^I_{ie} \gamma \left[\left(1 - \kappa_{g(i)} - \delta + \kappa_{g(i)} \delta\right) \left(\overline{w}^I \chi_e - r_i \overline{h} \right) + \left(\kappa_{g(i)} + \delta - \kappa_{g(i)} \delta \right) \underline{w} \right] + \\ L_i r_i \overline{h} + (1 - \mu) Y_i = r_i H_i^{g(i)}. \end{aligned} \tag{16}$$

Equilibrium in our model depends on a set of city-wide parameters, as well as parameters specific to home locations, work locations, or both. These parameters determine three sets of endogenous variables, as follows.

Equilibrium. *Given parameters* $\left\{ \overline{\boldsymbol{L}}, \boldsymbol{\chi}, \eta, \theta, \underline{w}, \overline{w}^I, \delta, \gamma, \overline{h}, \nu, \boldsymbol{\psi}, \boldsymbol{\kappa}, \rho_B, \Sigma, \overline{\pi}, \Xi_0, \Xi_1, \sigma, \ p(\cdot), \mu, \xi, \rho_Z, \epsilon \right\}$ *and model geography* $\left\{ \mathscr{F}, \mathscr{S}, \boldsymbol{g}(\cdot,\cdot), \boldsymbol{\iota}, \boldsymbol{\varsigma}, \boldsymbol{\tau}, \boldsymbol{\lambda}, \overline{\boldsymbol{T}}, \overline{\boldsymbol{Z}}, \boldsymbol{\Delta} \right\}$*, an equilibrium is a vector of location built heights* $v_i^{g(i)}$*, a vector of formal wages* w_{je} *by location and skill, and a vector of location floorspace rents* r_i *satisfying (i) the developer first-order condition* (8)-(9) *by neighbourhood type; (ii) labour market clearing* (15)*; and (iii) floorspace market clearing* (16)*.*

Territorial Informality in the Model

Our model captures the key dimensions implied by territorial informality discussed in Section 3. First, *crime exposure* plays an important role in the model because of individuals' risk aversion.

Hence, higher crime rates in slums ($\kappa_S > \kappa_F$) constitute a strong disamenity, particularly to higher-income residents who experience a higher loss if victimized.

Second, *slums are not subject to formal land regulation*, and therefore slum developers can build a factor σ more floorspace per unit land. However, they face a risk of expropriation that increases with height built due to the lack of property rights and legal enforcement of contracts. Thus, slums will have shorter buildings but can fit more residents and firms into a given building height. Whether slums have a greater total supply of housing and thus lower rents, versus similarly-located formal neighborhoods, depends on which of these two forces dominates.

Third, the model incorporates the distribution of *infrastructure and schools* observed in the data. Thus, low (high) access to infrastructure and school quality translates into lower (higher) exogenous amenities, $\overline{B}_i = \mathrm{g}_i(\iota_i, \varsigma_i)$. The residual component inherent in g_i is backed out from observed location population, thus capturing other unobserved disamenities of slums.

Fourth, our spatially-granular model of production differentiates between formal and informal employment and allows for substantial spatial variation in productivity and wages. The model is therefore able to capture a key consequence of *slums' embeddedness* near central, productive locations—namely, the fact that slum residents often have high *commuter market access* (see Section 3). Moreover, our formal job offer framework allow us to capture extensive margin barriers faced by residents of less-conveniently-located neighborhoods.

Thus, the model captures the key trade-off faced by slum residents: high access to formal jobs (and schools) at potentially low rents, but at the cost of high crime risk and potentially worse amenities. How low-skill residents and indeed aggregate output would benefit from formalization of slums along these different dimensions depends crucially on equilibrium responses. Reductions in crime risk or improvements in infrastructure, for example, might naturally raise rents and gentrify slums, a process strengthened by amenities spillovers. At the same time, lower slum crime could allow more workers to live in slums, commute to productive jobs, leverage agglomeration externalities, and thus boost aggregate output. In contrast, elimination of developers' expropriation risk, on its own, could actually lower slum rents. In the remainder of the paper, we discuss model estimation and how we use it to quantify the effects of these different consequences of territorial informality in general equilibrium.

5 Calibration

We calibrate our model using our novel highly spatially-granular data discussed in Section 2.2. To reduce the computational burden, we restrict our estimation and counterfactual exercises to the metropolitan area of Rio de Janeiro. Rio makes for a particularly instructive case study because it contains a large number of neighborhoods with comparable characteristics to the full sample of Brazilian urban neighborhoods discussed in Section 3. Our set of locations in the model

consists of 624 "area ponderacao" (henceforth AP), which roughly correspond to U.S. census tracts. APs typically correspond to neighborhoods, but large and populous neighborhoods are comprised by more than one AP. Since an AP aggregates multiple census blocks, we use "AP $\times$ slums" as unit of analysis in our calibration. We define three skill levels: less than a high-school education ($e = L$), high school ($e = M$), and university-educated ($e = H$).

We first specify key functions in our model. Second, we externally calibrate a subset of parameters based on aggregate and neighborhood-level data as well as established estimates in the literature. Third, we internally calibrate key human capital, preference, and housing supply parameters to match moments in our combined datasets and invert the model to obtain home-location-specific exogenous amenities and work-location-specific formal-sector productivities. Table 4 summarizes our calibration strategy.

Model Parameterization

We now parameterize the functions in our model. We parameterize formal commute costs as a function of straight-line distance Δ_{ij} between home and work locations, such that $\tau_{ij} \equiv \exp(\rho\Delta_{ij})$, where ρ captures the disutility of commute distance. We assume a linear specification for the exogenous amenity function $\mathrm{g}(\cdot,\cdot)$, whereby $\overline{B}_i \equiv \beta_i^0 + \beta^\iota \iota_i + \beta^\varsigma \log \varsigma_i$. The intercept β_i^0 captures location-specific residual, unobserved amenities; the β^c coefficients capture the effect of characteristic c on exogenous amenities, either infrastructure ι_i or school quality ς_i. Finally, expropriation risk satisfies $\mathrm{p}\left(v_i^S\right) \equiv \alpha \exp\left(v_i^S - 1\right)$. Higher values of the parameter α correspond to a steeper expropriation risk, i.e. one which rises more strongly with building height.[11]

External Calibration

We externally calibrate a series of preference, labour market, and production parameters based on the literature and data. We take the home and workplace Frechet shapes $\eta = 2.16$ and $\theta = 3.398$, respectively, from Tsivanidis (2023), where the latter equals the average of his values for low- and high-skill workers. We set the housing Cobb-Douglas share γ to the mean housing expenditure share of the top 25% of households by income in the POF expenditure data. Relative risk aversion, as in Lagakos et al. (2023), equals $\nu = 2$. On the production side, we set formal firms' Cobb-Douglas labour, $\mu = 0.794$, and the CES elasticity of substitution over skill groups, $\xi = 1.43$, as in Tsivanidis (2023). The informal job loss probability $\delta = 0.084$ equals

[11] Additionally, to avoid negative consumption values and for logical consistency, we assume that workers receiving a nonzero formal or informal wage but whose income, net of the cost $r_i\overline{h}$ of the subsistence requirement, is less than $\underline{w}$ receive the same total government benefits, $\underline{w}$ plus the cost of the subsistence requirement, as crime or job-loss victims. In this corner-solution-type case, however, we assume that formally-employed workers continue to supply their human capital χ_e to the formal firm, which pays the standard total labour cost, $w_{je}\chi_e$, to an absentee government.

Table 4: Calibration Strategy

Parameter	Description	Method	Data/Source
Preferences			
η	Home Frechet Shape	External	Tsivanidis (2023)
θ	Formal Work Frechet Shape	External	Tsivanidis (2023)
ν	Relative Risk Aversion	External	Lagakos et al. (2023)
γ	Housing Share	External	POF
ψ_L	Low-Skill Amenity Preference	Internal: relative median slum high-school share	Census (2000)
Labour Market			
ρ	Commute Distance Disutility	Internal: commute gravity	Chaves et al. (2023)
$\overline{w}^I$	Informal Wage	Internal: ratio of median informal to formal wage	Census (2000)
χ_M, χ_H	Human Capital by Skill	Internal: skill premia	Census (2000)
δ	Informal Pr(Job Loss)	External	Ulyssea (2010)
$\underline{w}$	Government Benefit	External	POF
Production			
μ	Labour Share	External	Tsivanidis (2023)
ξ	Skill Elasticity of Substitution	External	Tsivanidis (2023)
Housing Supply and Neighbourhood Characteristics			
$\overline{h}$	Housing Subsistence Requirement	Internal: bottom-20% housing share	POF
σ	Slum Floorspace Factor	Internal: relative median slum rent	PNAD
α	Expropriation Risk Parameter	Internal: slum housing supply elasticity	Google Open Buildings + Cadunico
Ξ_0	Height Cost Intercept	Internal: median building height	Google Open Buildings
Ξ_1	Height Cost Convexity	External	Google Open Buildings + Cadunico
$\overline{\pi}$	Land Coverage	External	Google Open Buildings
$\boldsymbol{\kappa_g}$	Pr(Crime Victimization)	External	SUS
$\overline{L}_e$	Population by Skill	External	Census (2000)
Geography			
$\boldsymbol{\beta^0}, \beta^\iota, \beta^\varsigma$	Exogenous Amenity Function	Model Inversion: population	Census (2000)
$\overline{\boldsymbol{Z}}$	Exogenous Formal Productivities	Model Inversion: formal employment by workplace	RAIS
$\boldsymbol{\lambda}$	Pr(Formal Job Offer)	External	Census (2000)
$\boldsymbol{\iota}, \boldsymbol{\varsigma}$	Infrastructure, School Quality	External	Census (2010), IDEB
$\overline{\boldsymbol{T}}$	Area	External	geobr
$\boldsymbol{\Delta}$	Distances	External	geobr
Spillovers			
ϵ	Agglomeration Elasticity	External	Ahlfeldt et al. (2015)
Σ	Amenity-High-Skill Elasticity	External	Gechter and Tsivanidis (2023)
ρ_B, ρ_Z	Spillover Decay Parameters	External	Set equal to $\hat{\rho}$.

Notes: Table lists model parameters along with calibration method and data or source used.

the yearly transition rate from informal employment to unemployment from Ulyssea (2010). To obtain an estimate of the income losses from crime and job loss, we set the government benefit $\underline{w} = 1$ such that the ratio between between $\underline{w}$ and the mean income predicted our model approximately equals the ratio between the average government benefit and average household income, conditional on receipt of nonzero government benefits, observed in the POF.

The remaining externally-calibrated parameters govern floorspace supply, neighbourhood characteristics, and spillovers. The height cost convexity $\Xi_1 = 1.44$ is the reciprocal of the formal housing supply elasticity, which we estimate directly using Google Open Buildings heights and Cadunico rents. The fixed land coverage $\overline{\pi} = 0.37$ comes from the mean ratio of built-up to total area across AP $\times$ slum units in our Google Open Buildings data. Crime risk by neighborhood type $\boldsymbol{\kappa_g}$ equals the population-weighted mean across either slum or formal blocks of the number of hospitalizations by firearm-related trauma per capita.[12] We take populations $\overline{L}_e$, separately by our three skill levels, as the number of individuals aged 25 or over in the Rio de Janeiro metro area in the 2000 Census. As for spillovers, the agglomeration elasticity equals Ahlfeldt et al. (2015)'s estimate of $\epsilon = 0.07$. We set the elasticity of amenities with respect to nearby high-skill density equal to Gechter and Tsivanidis (2023)'s estimate of $\Sigma = 0.88$ from India. Finally, we set the spillover decay parameters equal to the distance semi-elasticity of commute costs, ρ, which we internally calibrate in the next step, such that $\rho_B = \rho_Z = \hat{\rho}$.

Lastly, we calibrate the model geography. We calculate straight-line distances Δ_{ij} between home and work location centroids using shapefiles provided by the geobr package. The formal offer probabilities λ_i equal the share of workers who are not informally- or self-employed in a given AP $\times$ slum in the 2000 Census. Location area $\overline{T}_i$ also equals AP $\times$ slum area from geobr shapefiles. We measure infrastructure ι_i as the simple average of the shares of households in a given AP $\times$ slum who report piped water access, sewer network access, garbage collection, street lights, paved roads, street trees, sidewalks, no trash on the streets, and no open sewers around their homes. Finally, to measure school quality ς_i, we take an inverse-distance-weighted average of AP $\times$ slums' average IDEB Education Index[13]

Internal Calibration and Model Inversion

We solve the full model to match nine auxiliary models or moments and thereby calibrate the remaining labour market, preference, and housing supply parameters. Conditional on these parameters, we follow in the tradition of Ahlfeldt et al. (2015) and invert the model to obtain home-location exogenous amenities and work-location formal productivities.

[12] Since our hospitalization rates are measured over the course of 4 years, we scale by $70/4$ to obtain lifetime rates, given the year-2000 average life expectancy in Brazil of 70 years, as reported by the World Bank.

[13] Given straight-line distance Δ_{ij}, the weights equal $1/\exp(0.5 \times \Delta_{ij})$.

Auxiliary Models For our first auxiliary moment, we employ mobile phone data published by Chaves et al. (2023) to estimate a commute gravity regression. For an origin-destination geography which corresponds to subdistricts of Rio municipality plus other municipalities in the metropolitan area, we approximate commute flows using their aggregations of mobile phone pings. Specifically, we measure the probability $\pi_{j|i}$ that a resident of location i appears most frequently during working hours on workdays in a given neighborhood j. Our auxiliary model then consists of a simple regression of these workplace choice probabilities on straight-line origin-destination distance, $\log \pi_{j|i} = \beta_0^1 + \beta_\Delta^1 \Delta_{ij} + \delta_i + \delta_j + \varepsilon_{ij}$, with origin-destination fixed effects and where ε_{ij} denotes the idiosyncratic error.

Our next four auxiliary moments draw on the entire population in Brazilian metro areas in the 2000 Census. The second auxiliary moment equals the ratio of median informal to median formal wages in metro areas nationwide in the 2000 Census. The third and fourth moments draw on the same Census sample and capture skill premia for high-school and university education. In the former case, we take the ratio of the mean income of employed workers with a high-school but not university education to the mean income of those without a high-school education; in the latter case, we take the ratio of the mean income of workers with a university education to the mean income of those with high-school but not university. The fifth moment captures skill sorting into, or rather out of slums: the population-weighted median share of slum residents aged 25 or over with a high-school education, relative to the corresponding median share in formal neighbourhoods.

Our sixth and seventh auxiliary models employ household survey data. In the nationwide POF expenditure data, we calculate the mean housing expenditure share of renters in bottom 20% of households by income. In a panel of the PNAD survey over the period 2001-2009, we calculate the median reported rent per room in slums, again relative to the median in formal neighbourhoods nationwide.

Our final two auxiliary models use our Google Open Buildings database on building footprints and heights across all of Brazil's metro areas. The eighth moment measures overall median building height.[14] For the ninth moment, we additionally draw on AP $\times$ slum average rent per room r_i as recorded in the Cadunico to estimate a slum housing supply elasticity. Specifically, we regress the log of the aforementioned mean building height $v_i^{g(i)}$ on rents according to $\log v_i^{g(i)} = \beta_0^9 + \beta_r^9 \log r_i + \epsilon_i$.

Identification We now provide the usual heuristic argument for identification of the nine parameters estimated via internal calibration; Table 5 lists parameters and corresponding moments in both the data and model.

[14]Specifically, we calculate an area-weighted mean building height in metres and then normalize by the average floor-to-floor height of 3.2 metres, as reported by a London planning study.

Table 5: Internal Calibration

Moment			Parameter		
Description	*Data*	*Model*	*Description*		*Value*
Commute Gravity: Distance Coefficient	-0.058	-0.077	ρ	Commute Distance Disutility	0.035
Median Informal/Formal Wage Ratio	0.89	—	$\overline{w}^I$	Informal Wage	25
Skill Premium: High-School	2.06	2.08	χ_M	High-School Human Capital	1.6
Skill Premium: University	2.89	2.94	χ_H	University Human Capital	6.5
Relative Slum High-School Share	0.36	—	ψ_L	Low-Skill Amenity Preference	0.15
Bottom-20% Housing Share	0.28	0.27	$\bar{l}$	Housing Subsistence Requirement	105
Relative Slum Rent	0.79	0.70	σ	Slum Floorspace Factor	1.6
Median Building Height	1.6	—	Ξ_0	Height Cost Intercept	0.007
Housing Supply Elasticity	0.337	—	α	Expropriation Risk Parameter	0.06

Notes: This table displays the nine auxiliary moments used in internal calibration, listing the values in the data and in the model, using which we simulate 200,000 workers and run equivalent specifications. I also list the model parameter heuristically corresponding to each moment, along with its internally-calibrated value.

We begin with labour-market and human capital parameters. The commute gravity regression speaks to the disutility ρ of commute distance. Though this "usual" commute gravity regression does not hold exactly in our model, the sensitivity β^1_Δ of location choices to commute distance nonetheless increases with ρ. Next, the ratio of mean formal to informal wages, though also a function of home and work location choices, varies strongly with informal wage $\overline{w}^I$. The high-school and university skill premia pin down the magnitude of χ_M and χ_H, respectively, relative to the low-skill productivity χ_L, which we normalize to 1.

We now move to the internally-calibrated preference parameters. Slums' relative high-school-educated share share pins down the amenity preference of low-skill workers, ψ_L. Intuitively, we infer that slums have much worse amenities, so the less the low-skill value amenities, i.e. the lower ψ_L, the lower the relative high-school share in slums. We focus on this non-high-school versus high-school margin of skill sorting and thus set $\psi_M = \psi_H = 1$. The bottom-20% housing expenditure share identifies the housing subsistence requirement $\overline{h}$, as the latter, by construction, pushes up the housing spending of the poorest households most acutely.

The last parameters to be calibrated govern housing supply. The relative rent in slums versus formal neighbourhoods identifies the slum floorspace factor σ, which decreases slum rents through a direct supply-shift effect. The median building height identifies the height cost intercept Ξ_0: higher Ξ_0 raises the marginal cost of height and thus lowers equilibrium building heights. Finally, the estimated housing supply elasticity β^9_r varies inversely with the expropriation risk parameter α, since higher expropriation risk constrains developers' height choices and thus reduces the responsiveness of housing supply to rents.

Model Inversion Conditional on all preceding parameters, we then invert the model to recover the rest of the geography. Namely, we match population by home location, i.e. AP $\times$ slum, from the 2000 Census to recover composite home location amenities B_i. To obtain work locations' composite formal productivities Z_j, we exactly match formal employment by workplace, as observed in RAIS. We then substitute model-predicted high-skill density $L_{iH}/\overline{T}_i$ as well as the externally-calibrated spillover parameters into Equation (13) to recover the exogenous component $\overline{B}_i$ of amenities. Similarly, we substitute model-predicted employment E_j into Equation (14) to recover the exogenous productivity "intercept" $\overline{Z}_j$. In a final step, we regress the estimated exogenous amenities $\overline{B}_i$ on measured infrastructure ι_i, log school quality ς_i, and a slum dummy to recover the corresponding coefficients β^ι and β^ς in our parameterized amenity function. To obtain residual neighbourhood amenities β_i^0, we use these coefficients to net out infrastructure and school quality effect from the calibrated total exogenous amenities $\overline{B}_i$.

6 Slums in the Model

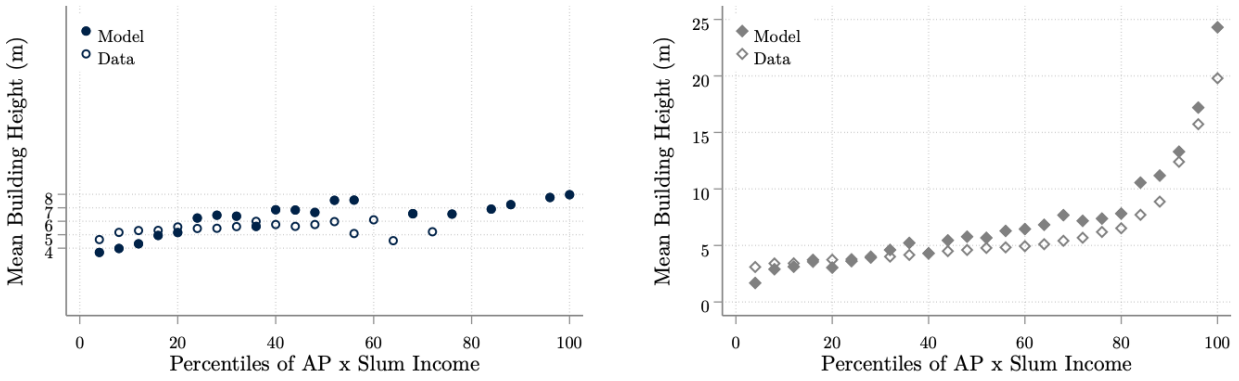
We now assess whether the model can replicate key features of slums' unique bundle: constrained vertical development, embeddedness in desirable city locations, inferior infrastructure, and superior access to formal jobs and good public schools. In particular, we return to the stylized facts from Section 3 and compare the corresponding model-implied objects for Rio de Janeiro with their empirical counterparts. Importantly, we consider only moments not directly targeted in the calibration.

Fact 2 emphasized building height as a distinctive feature of slums. Figure 10 shows that the model replicates the empirical height gradients plotted for slums and formal neighbourhoods in Rio de Janeiro. Most importantly, our developer expropriation probability prevents slum height from rising steeply at the upper end of the neighbourhood income distribution, as occurs in desirable, high-income formal areas.

Next, Fact 4 showed that slums "embed" themselves in similar locations as rich formal neighbourhoods. Figure 11 plots the empirical distributions of distance segregation by neighbourhood type — slum, poor formal, and rich formal — in Rio, alongside the same distributions in our baseline model equilibrium. The endogenous adjustment of productivities, wages, and rents, combined with the locational fundamentals our model infers, generate exactly the same pattern observed in the data: slums and rich formal neighbourhoods tend to occupy desirable, low-distance-segregation locations, with poor formal neighbourhoods shunted to the periphery.

Finally, Figure 12 assesses whether the model reproduces the infrastructure and access enjoyed by neighbourhoods at different income levels, as documented in Facts 1 and 5. The top two panels display the infrastructure index ι_i versus percentiles of neighbourhood income in the data and model, separately for slums on the left in Panel (a) and formal neighbourhoods on the right

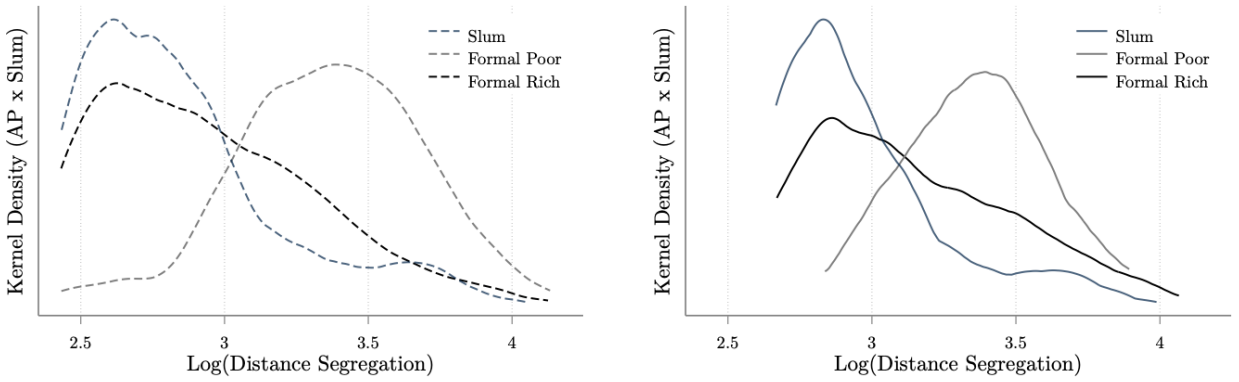
Figure 10: Model-Implied Building Height vs. Neighbourhood Income



(a) Slum (b) Formal

Notes: Panels (a) and (b) display binned scatterplots of model-implied building height $v_i^{g(i)}$ against percentiles of mean neighborhood income across AP × slum locations in our baseline model equilibrium and across tracts in Rio de Janeiro in our data, for slums and formal neighbourhoods, respectively.

Figure 11: Slum Embeddedness



(a) Data (b) Model

Notes: Panel (a) displays the distribution of log distance segregation by neighbourhood type across tracts in Rio de Janeiro our data, corresponding to Fact 4, where we define distance segregation as the average distance between a location and locations with mean resident income above the 75th percentile of the city income distribution. Rich formal neighbourhoods are those with mean income above the across-location median. Panel (b) displays the corresponding distribution across AP × slum locations our baseline model equilibrium.

in Panel (b). As in the data, slum infrastructure access falls uniformly below that of formal neighbourhoods across the income distribution, and the gap grows at higher neighbourhood incomes.

The middle two panels turn to commuter market access, an important advantage of slums. Figure 8 underscored the unique "quadrant" occupied by slums: low-income but close to dense clusters of formal jobs. Importantly, in our model, locations' formal job density, and thus

commuter market access, arises in equilibrium through the interaction of firms' labour demand, workers' commute choices, and agglomeration spillovers. These rich general-equilibrium forces generate high CMA in slum neighbourhoods at low income levels in Panel (c), just as in the data, while formal CMA in Panel (d) rises similarly steeply with neighbourhood income in data and model.

The bottom panels display our inverse-distance-weighted IDEB school quality measure. Recall that, in the data, slums benefit from above-average nearby school quality across all income levels in our model equilibrium, while the typical poor formal neighbourhood has access to only middling schools. The model reproduces this pattern — poor slums can nonetheless co-exist embedded in desirable city locations close to high-quality public schools.

7 Territorial Informality, Output, and Welfare

We have shown that slums' territorial informality gives rise to a unique bundle of neighborhood characteristics which offer sometimes stark trade-offs to individual residents; the evidence in Section 3 suggests that slums' existence may have important consequences for aggregate productivity and welfare. In this final section, we thus leverage our estimated model to quantify exactly these aggregate trade-offs. First, we quantify how the unique bundle slums offer translates into their residents welfare *vis-a-vis* that of formal areas, as well as how much each of slums' traits contribute to the overall gap relative to formal areas. Second, we simulate policy counterfactuals which address individual elements of the "slum bundle", and correspond to different classes of prototypical slum policies. Crucially, our model allows us to quantify not only how heterogeneous residents re-sort across neighborhoods, but also how these adjustments spill over to amenities, productivity, and ultimately welfare levels in the city.

Accounting for Slum Welfare

First, we investigate how the slum trade-offs induced by territorial informality add up: are slum dwellers, on average, worse off than formal-neighbourhood residents, and why? Since our Frechet idiosyncratic shocks have the well-known knife-edge property by which they equalize utility across locations, we restrict our attention to the neighbourhood average of the deterministic component of utility, given by the utility in Equation (1) net of the idiosyncratic home and workplace shocks, in equilibrium. The first bar in Figure 13 demonstrates that the average slum resident has deterministic utility 27% below their poor-formal-neighbourhood counterparts, measured in terms of equivalent variation in slum-dweller incomes.

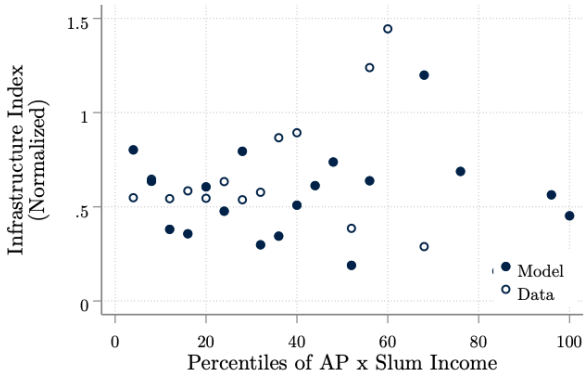
In the remaining bars in Figure 13, we decompose this slum utility differential into the contributions of the different pieces of the "slum bundle." In particular, starting from the raw

27% utility differential between slums and poor formal neighbourhoods, we equalize each slum characteristic to the formal equivalent one at a time, displaying the resulting slum-versus-poor-formal utility differential in each case. We begin with crime: equalizing slum crime rates to those of formal areas, $\kappa_S = \kappa_F$, reduces the gap to just 7%. Since workers are risk averse and slums suffer from almost twice the crime rate of formal areas, this crime reduction eliminates over three quarters of the slum penalty. The effect of slums' lack of property rights on housing supply is similarly consequential. Elimination of the risk of developer expropriation via $\alpha = 0$ nearly closes the slum-poor formal gap entirely; the resulting expansion in housing supply pushes slum rents sharply downward, to the benefit of all workers. In contrast, when we impose formal zoning constraints on slums, i.e. eliminate slums' additional floorspace and set $\sigma = 1$, the slum penalty widens to 36%, as the resulting supply contraction raises rents.

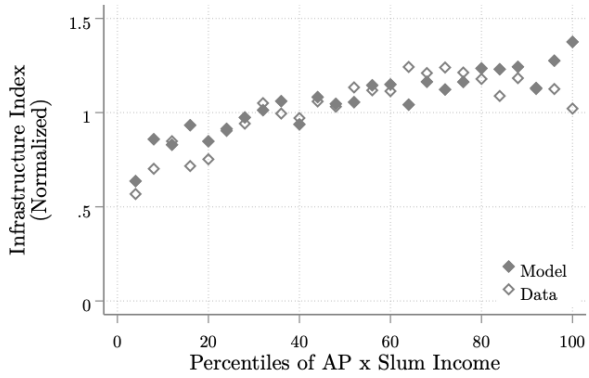
We next consider slums' amenity deficits. Equalization of slums' average infrastructure index to the formal counterpart reduces the slum utility penalty from 27% to 20%, as a result of our strong estimated link β^{ι} between infrastructure and amenities. Recall that slums actually enjoy superior access to high-quality schools, relative to the average formal neighbourhood; were slums to lose this advantage, the slum utility penalty would modestly narrow to 21%. In the final bar, we boost all slums' residual amenities β_i^0 by the corresponding difference in formal vs. slum means; the slum penalty falls to 14%. Each of the dimensions of territorial informality we consider can thus individually account for a substantial share of the slum penalty, with crime and the lack of property rights emerging as the most powerful channels. The results also reveal a clear trade-off: slum residents suffer higher crime, poor infrastructure, and inferior amenities, but the absence of zoning restrictions boosts housing supply, such that even conveniently-located slums can maintain lower rents.

Figure 12: Model-Implied Access and Infrastructure vs. Neighbourhood Income

Infrastructure Index

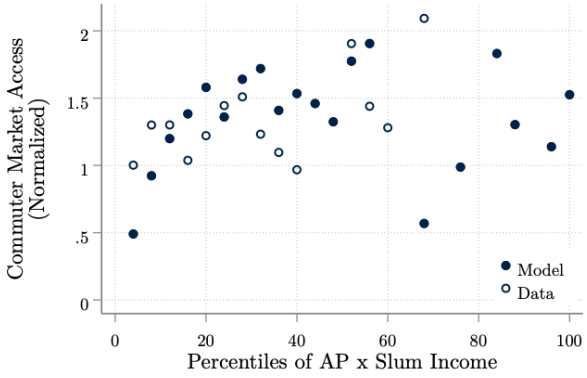


(a) Slum

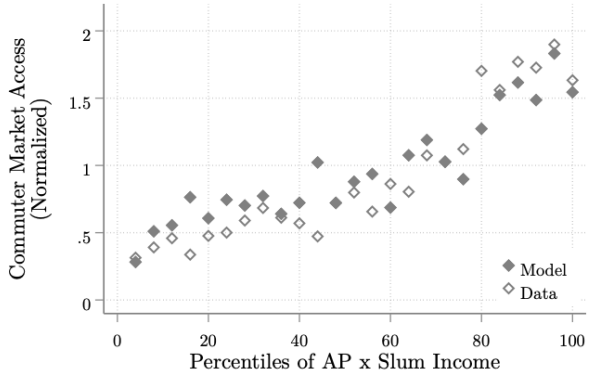


(b) Formal

Commuter Market Access

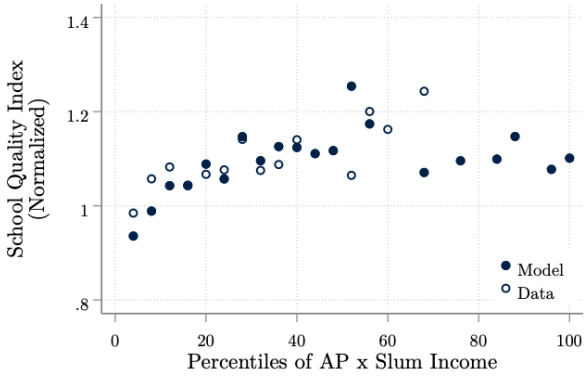


(c) Slum

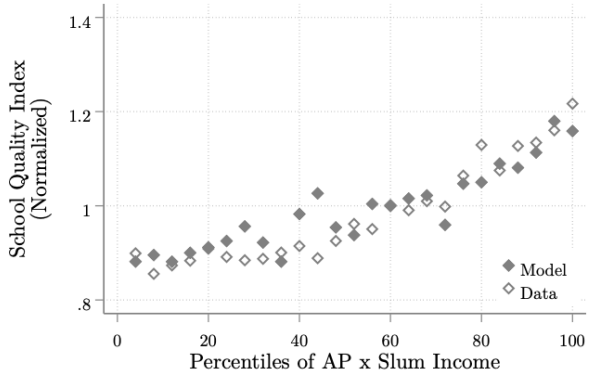


(d) Formal

Public Schools, Quality-Weighted



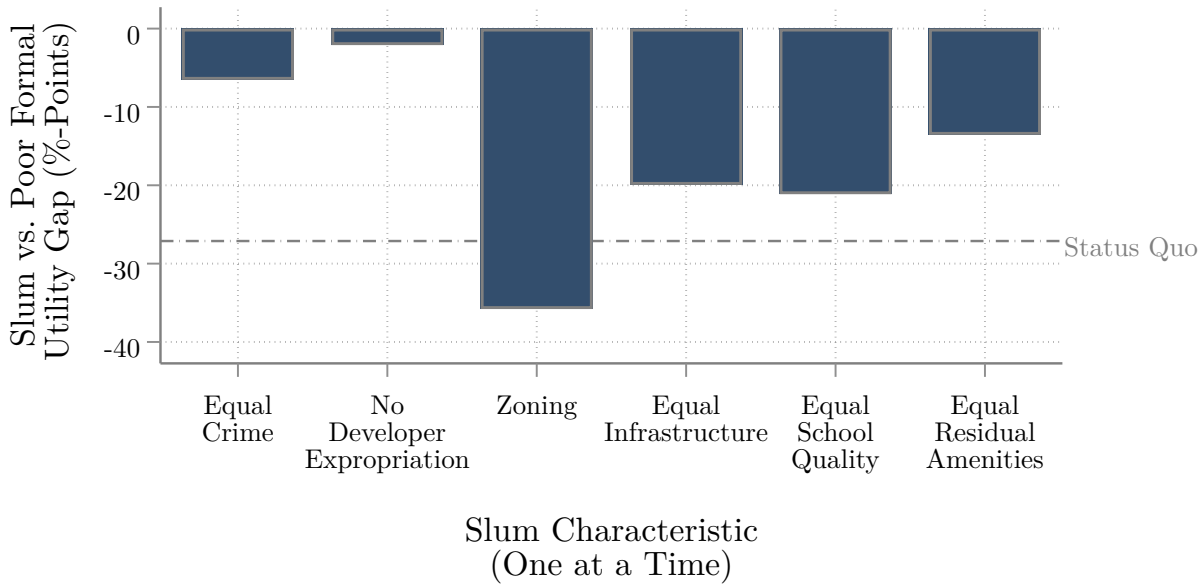
(e) Slum



(f) Formal

Notes: Panels (a) and (b) display binned scatterplots of the infrastructure index ι_i against (percentiles of) mean neighborhood income across AP $\times$ slum locations in our baseline model equilibrium and across tracts in Rio de Janeiro in our data, for slums and formal neighbourhoods, respectively. Panels (c) and (d) instead display equivalent plots for commuter market access, $\text{CMA}_i \equiv \sum_j E_j/\tau_{ij}^{\theta}$, on the vertical axis, while Panels (e) and (f) display inverse-distance-weighted school quality ς_i. Note that we normalize infrastructure, CMA, and school quality by dividing by the population-weighted average across percentiles.

Figure 13: Decomposition of Slum Utility Penalty



Notes: Figure displays the results of our decomposition of the deterministic utility of the average slum resident, relative to the average poor formal neighbourhood resident, in percentage terms. We measure deterministic utility as in Equation (1) but net of the idiosyncratic home and workplace shocks. The dotted line indicates the baseline slum vs. poor formal utility differential, calculated as equivalent variation in slum residents' incomes. Each bar displays the resulting slum vs. poor formal utility differential when we equalize each slum characteristic to that of formal neighbourhoods, one at a time. From left to right, we (i) equalize crime rates $\kappa_S = \kappa_F$, (ii) eliminate the expropriation risk $\alpha = 0$, (iii) introduce zoning, i.e. eliminate slums' additional floorspace factor via $\sigma = 1$, (iv) increase all slums' infrastructure ι_i by the difference in mean ι_i between slums and all formal neighbourhoods, (v) decrease slums' school quality ς_i by the corresponding difference in means, and (vi) increase slum amenities by the mean difference in exogenous amenities $\overline{B}_i$ between slums and all formal neighbourhoods, conditional on infrastructure and school quality.

Formalization Counterfactuals

We now turn to the welfare effects of prototypical formalization policies which correspond to each element of the slum bundle and represent different classes of policies often discussed in developing-world cities. For ease of interpretation, we initially implement each counterfactual one at a time, as in the decomposition exercise above. Panel A of Table 6 reports (i) the resulting changes in total formal plus informal output, $Y \equiv \sum_j Y_j^F + \sum_{i,e} L_{ie}^I \overline{w}^I \chi_e$; (ii) the aggregate welfare gains, measured in percent equivalent variation relative to the baseline equilibrium;and (iii) welfare inequality across skill groups, measured as low- or medium-skill welfare as a proportion of high-skill welfare, where we measure differences in terms of equivalent variation to high-skill incomes. We then implement the same counterfactuals cumulatively, in two different orders, in Panels B and C.

We first focus on the one-by-one counterfactuals in Panel A to cleanly isolate the effects of each policy. Consider, in the first row of Table 6, the crime reduction counterfactual, which equalizes slum crime rates to formal crime rates. Workers are sufficiently risk-averse that all skill groups benefit substantially from anti-crime measures in slums; aggregate welfare rises by 4.5%. These gains stem not only from the direct utility boost from less-frequent victimization but also from an over 4% increase in aggregate output, itself the largest of any of the policies we consider. Lower crime allows more workers to choose to live in slums, themselves often embedded near productive formal jobs. Exactly the high-income workers with the most to lose from crime also possess the highest human capital and thus stand to boost output the most strongly when they work in more productive locations.

The amenity-related counterfactuals go hand-in-hand with similarly substantial welfare gains. Infrastructural upgrades, which we implement as an increase in each slum's ι_i by the average formal-slum differential in infrastructure access, boost welfare by 4%, and equalization of (average) amenity residuals between slums and formal neighbourhoods increases welfare by almost 14%. Because we estimate that they value amenities more highly, medium- and high-skill workers gain even more, such that welfare inequality, as measured by relative low-skill welfare in the final two columns, nonetheless rises. As in the case of crime, more workers sort into slums and can thus access nearby formal jobs; as a result, output rises by 0.4% and 0.7%, respectively. Taken together, these large gains underscore the costs of the "downsides" of the status-quo slum bundle.

Formalization policies which target the housing sector particularly unequal welfare effects, in significant part due to amenity spillovers. The second row of Table 6 reports the effects of eliminating the risk of developer expropriation in slums, which permits developers to build as tall as they would in formal areas. The resulting expansion in housing supply halves relative slum rents and generates the second-largest welfare gains of any policy. Inequality between low-

and high-skill workers, however, increases, as the relative welfare of the former falls from 0.67 to 0.59. Due to the non-homotheticity in housing demand, low-skill workers benefit disproportionately from lower rents, but they also place a lower value on amenities — which rise due to inflows of high-skill workers. Worker re-sorting away from productive work locations actuall reduces aggregate output by 1.1%. Conversely, imposing formal zoning constraints on slums — eliminating their additional floorspace via $\sigma = 1$ — reduces welfare for all groups; the relative welfare measures in the last two columns indicate that, strikingly, high-skill workers now lose the most. The implicit supply constraint imposed by zoning raises slum rents, pushing workers out of slums; the resulting fall in high-skill density reduces amenities. Because high-skill workers value amenities the most, they bear the largest welfare loss. Nonetheless, as workers move to formal neighbourhoods, output rises by 1.5%.

Thus, each slum formalization policies introduces different trade-offs. In the case of three central downsides of territorial informality — crime, infrastructure, and amenities — cities can take comfort in the fact that formalization benefits all residents substantially, even if it does so unequally, and aggregate output rises in each case. Policies targeting the housing sector, however, present starker dilemmas. Removal of expropriation risk yields large welfare gains but modestly reduces aggregate output, as workers sort into less-productive jobs near slums. When it comes to enforcement of zoning regulations and other measures which effectively constrain housing supply, however, cities must tread carefully. Indeed, these policies, though not without their own, albeit unmodeled, benefits in terms of safety and health, aim directly at one of the key advantages of slums: their ability to accommodate large numbers of low-income residents at low cost.

Complementarities Most cities, of course, strive to eventually address all of these dimensions of territorial informality. How do the effects change when cities layer one policy on top of another, and does the order matter? To address these questions, we now implement the same formalization policies cumulatively, maintaining all preceding policies in place, and consider two different orderings. Panels B and C of Table 6 show the full results for each ordering; for ease of interpretation, we focus on low-skill workers. Figure 15(a) displays the low-skill welfare gains when property rights reforms — no developer expropriation and zoning — come first. Figure 15(b) instead shows the case where cities grant property rights last. The top of each bar displays the cumulative welfare gain from a given counterfactual implemented together with all preceding policies, from left to right, and the upper dotted line indicates the cumulative effect of the full formalization bundle — namely, an 18% welfare increase for low-skill workers, relative to the baseline equilibrium.

Importantly, the comparison between panels Figure 15(a) and Figure 15(b) reveals that the order of reforms matters substantially. In particular, we find that low-skill workers benefit from

Table 6: Formalization Counterfactuals

Counterfactual	% Δ Output	% Δ Welfare	Welfare Inequality	
			Low / High	Med. / High
Baseline	—	—	0.67	0.74
Panel A: One At A Time				
Equal Crime	4.15	4.54	0.67	0.78
No Developer Expropriation	−1.06	7.59	0.59	0.79
Zoning	1.52	−7.28	0.80	0.78
Equal Infrastructure	0.38	4.07	0.65	0.74
Equal Residual Amenities	0.69	13.66	0.55	0.75
Panel B: Cumulative, Property Rights First				
No Developer Expropriation	−1.06	7.59	0.59	0.79
+ Zoning	−0.44	3.47	0.61	0.75
+ Equal Crime	1.85	14.99	0.47	0.68
+ Equal Infrastructure	2.35	21.02	0.45	0.67
+ Equal Residual Amenities	3.24	38.14	0.31	0.67
Panel C: Cumulative, Property Rights Last				
Equal Crime	4.15	4.54	0.67	0.78
+ Equal Infrastructure	4.58	9.48	0.65	0.79
+ Equal Residual Amenities	5.22	26.71	0.43	0.79
+ No Developer Expropriation	3.15	47.93	0.31	0.73
+ Zoning	3.24	38.14	0.31	0.67

Notes: Table reports the percent change in output, aggregate welfare gains, and welfare inequality across skill groups from our formalization counterfactuals. Output measures total formal and informal output. We measure welfare changes in percent equivalent variation relative to the baseline equilibrium. The welfare inequality columns report low- or medium-skill welfare as a proportion of high-skill welfare, where we measure differences in terms of equivalent variation to high-skill incomes. Panel A implements each counterfactual one at a time. Panels B and C implement the same counterfactuals cumulatively in two different orders: Panel B introduces property rights, i.e. no developer expropriation and zoning, first, while Panel C introduces them last. The counterfactuals are: (i) equalizing crime rates $\kappa_S = \kappa_F$, (ii) eliminating the expropriation risk $\alpha = 0$, (iii) introducing zoning, i.e. eliminating slums' additional floorspace factor via $\sigma = 1$, (iv) increasing all slums' infrastructure ι_i by the difference in mean ι_i between slums and all formal neighbourhoods, and (v) increasing slum amenities by the mean difference in exogenous amenities $\overline{B}_i$ between slums and all formal neighbourhoods, conditional on infrastructure and school quality.

40 percentage-point larger welfare gains from the combined crime, infrastructure, and amenity upgrades when implemented after the alleviation of developer expropriation risk, as in Figure 15(a). This stark divergence again stems directly from non-homothetic housing demand. When cities improve slums after removal of the implicit height constraints imposed by a lack of property rights, low-skill workers can reap the benefits because rents rise only mildly. In contrast, crime and amenity improvements under more- constrained housing supply, in Figure 15(b), simply

gentrify the low-skill straight out of slums.

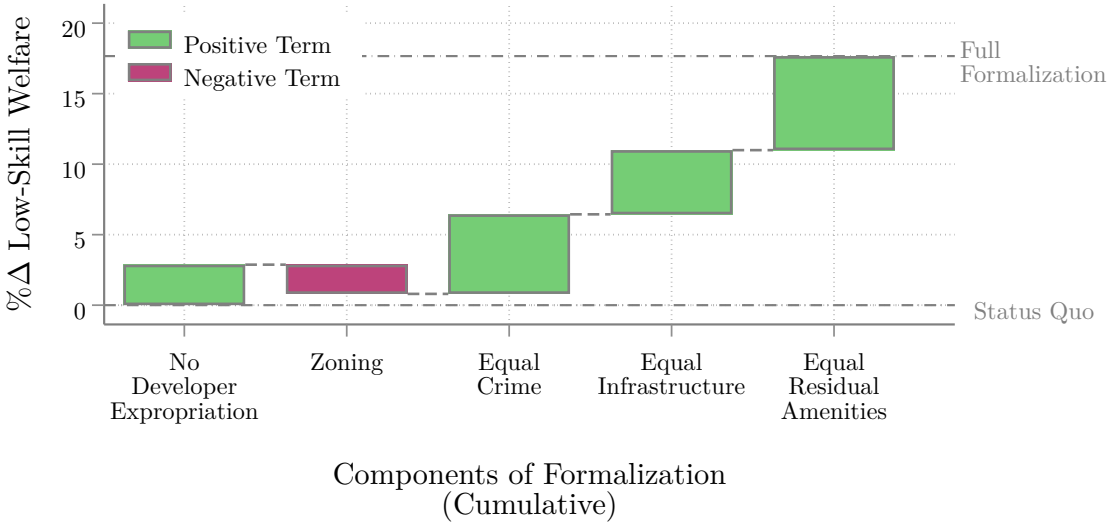
By a similar token, the zoning reform hurts low-skill workers more acutely in Figure 15(b), when implemented after crime-fighting and amenity upgrades. In this case, crime reductions and amenity improvements have already made slums attractive to the large masses of medium- and high-skill workers. As a result, the effective supply constraint introduced by zoning drives rents up more strongly and thus harms low-skill workers, who already spend the most on housing.

8 Conclusion

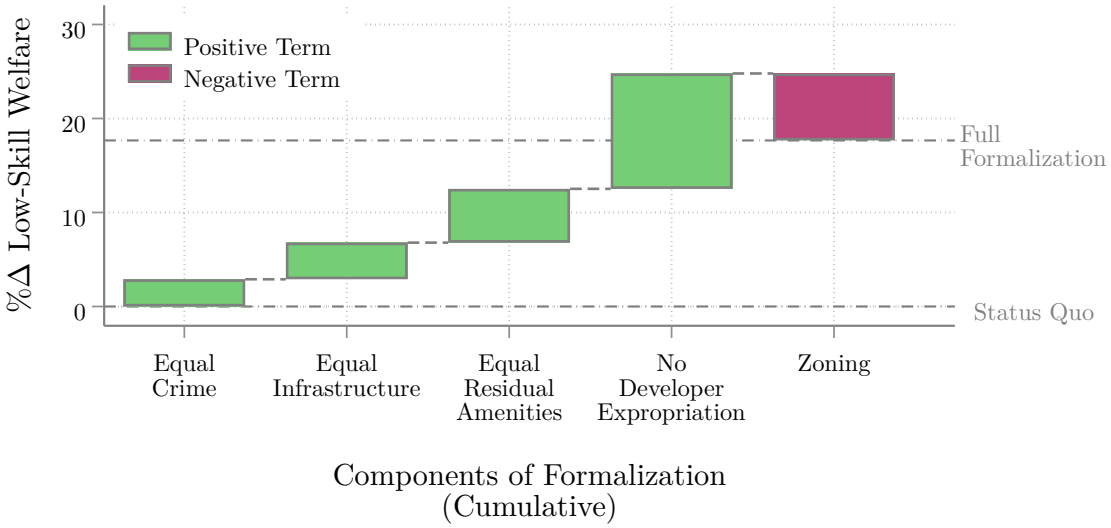
In this paper, we emphasize *territorial informality* as the key defining trait of slums. We leverage granular data sources covering nearly 80,000 census blocks across all of urban Brazil to show that slums can be conceptualized as "state voids"—areas characterized not by poverty or inadequate infrastructure, but fundamentally by weak state presence and authority—which allows them to offer a bundle of characteristics quite distinct from poor but formal areas. In particular, slums suffer from constrained building height, violence, and sometimes-poor infrastructure, but their "embedded" locations within the city proper provide superior access to formal jobs and good schools. We then develop a spatial general equilibrium model with heterogeneous, risk-averse workers, formal and informal housing markets, and endogenous amenity and productivity spillovers that rationalizes key consequences of territorial informality. We use the estimated model to quantify the relative importance of the different components of the slum bundle of characteristics and to quantify the city-wide tradeoffs associated with them.

Slum formalization policies turn out to introduce sharp trade-offs: aggregate output virtually always increases, often at the cost of greater inequality. Upgrades targeted at crime, infrastructure, and amenities nonetheless generate across-the-board gains. In contrast, reforms, such as zoning, which effectively constrain housing supply threaten slums' key advantage: their ability to accommodate large numbers of low-income residents at low cost. Cities must thus tread lest they lose the benefits that territorial informality confers on their most-disadvantaged residents.

Figure 14: Cumulative Formalization and Low-Skill Welfare



(a) Property Rights First



(b) Property Rights Last

Notes: Figure displays the low-skill welfare gains from our formalization counterfactuals, implemented cumulatively, always measured in percent equivalent variation relative to the baseline equilibrium. Panel (a) shows welfare changes when property rights reforms, i.e. no developer expropriation and no crowding, are implemented first, followed by the other counterfactuals in order from left to right. The lower dotted line indicates the baseline starting point, and the upper dotted line indicates the cumulative effect of all formalization counterfactuals. Each bar displays the cumulative welfare effect from a given counterfactual implemented together with all preceding it. From left to right, we (i) eliminate the expropriation risk $\alpha = 0$, (ii) introduce zoning, i.e. eliminate slums' additional floorspace factor via $\sigma = 1$, (iii) equalize crime rates $\kappa_S = \kappa_F$, (iv) increase all slums' infrastructure ι_i by the difference in mean ι_i between slums and all formal neighbourhoods, and (v) increase slum amenities by the mean difference in exogenous amenities $\overline{B}_i$ between slums and all formal neighbourhoods, conditional on infrastructure and school quality. Panel (b) shows the case where property rights reforms are instead implemented last.

References

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015). The economics of density: Evidence from the berlin wall. *Econometrica 83*(6), 2127–2189.

Alves, G. (2021). Slum growth in brazilian cities. *Journal of Urban Economics 122*.

Arias, E. D. and N. Barnes (2017). Crime and plural orders in rio de janeiro, brazil. *Current Sociology 65*(3), 448–465.

Asher, S., K. Jha, P. Novosad, A. Adukia, and B. Tan (2026). Residential segregation and unequal access to local public services in india: Evidence from 1.5 m neighborhoods. Technical report, National Bureau of Economic Research.

Balboni, C., G. Bryan, M. Morten, C. O'Connor, and B. Siddiqi (2025). Impacts of infrastructure: People versus place.

Belchior, C., G. Gonzaga, and G. Ulyssea (2025). Who benefits from social housing? experimental evidence from a large-scale program in brazil. Mimeo.

Blattman, C., G. Duncan, B. Lessing, and S. Tobon (2025). Statebuilding in the city: An experiment in civilian alternatives to policing. Accepted at American Political Science Review.

Cavalcanti, T., D. Da Mata, and M. Santos (2019). On the determinants of slum formation. *The Economic Journal 129*(621), 1971–1991.

Chaves, J. C., M. A. da Silva, R. de Souza Alencar, A. G. Evsukoff, and V. da Fonseca Vieira (2023). Human mobility and socioeconomic datasets of the rio de janeiro metropolitan area. *Data in Brief 51*, 109695.

Collier, P. and A. J. Venables (2014). Housing and urbanization in africa: Unleashing a formal market process. *World Bank Policy Research Working Paper* (6871).

De Souza, J., R. Willadino, R. Nascimento, F. Rodriguez, and F. L. Fernandes (2009). Escape routes: the path of young people in the social network of drug trafficking: journeys.

De Souza e Silva, J., F. L. Fernandes, R. Willadino Braga, and J. Global (2008). Grupos criminosos armados com domínio de território. *Segurança, tráfico e milícia no Rio de Janeiro. Rio de Janeiro: Fundação Heinrich Böll*, 16–24.

Deffebach, P., D. Lagakos, Y. Miyauchi, and E. Yamada (2025). The spatial distribution of income in cities: New global evidence and theory. Mimeo.

Dowdney, L. (2003). *Crianças do tráfico: um estudo de caso de crianças em violência armada organizada no Rio de Janeiro*. 7letras.

Ferreira, P. C., A. Monge-Naranjo, and L. T. de Mello Pereira (2024). Of cities and slums. *Forthcoming, Journal of Political Economy*.

Galiani, S. and E. Schargrodsky (2010). Property rights for the poor: Effects of land titling. *Journal of Public Economics 94*(9), 700–729.

Gechter, M. and N. Tsivanidis (2023). Spatial spillovers from high-rise developments: Evidence from the mumbai mills. *Revise & Resubmit accepted, Econometrica*.

Gertler, P., M. Gonzalez-Navarro, R. Undurraga, and J. A. Urrego (2025). In-situ upgrading or population relocation? direct impacts and spatial spillovers of slum renewal policies. Technical report, National Bureau of Economic Research.

Gonzalez-Navarro, M. and R. Undurraga (2023). Immigration and slums.

Gulyani, S. and D. Talukdar (2008). Slum real estate: The low-quality high-price puzzle in nairobi's slum rental market and its implications for theory and practice. *World Development 36*(10), 1916–1937.

Harari, M. (2024). Residential patterns and local public goods in urban brazil.

Harari, M. and M. Wong (2024). Slum upgrading and long-run urban development: Evidence from indonesia. *Accepted, Review of Economic Studies*.

Heblich, S., S. J. Redding, and D. M. Sturm (2020). The making of the modern metropolis: Evidence from london. *The Quarterly Journal of Economics*, 2059–2133.

Henderson, J. V., T. Regan, and A. J. Venables (2021). Building the city: From slums to a modern metropolis. *Review of Economic Studies 88*, 1157–1192.

Henderson, J. V., A. J. Venables, T. Regan, and I. Samsonov (2016). Building functional cities. *Science 352*(6288), 946–947.

IBGE (2024). Favelas e comunidades urbanas. Notas metodologicas n. 01 – Sobre a mudança de Aglomerados Subnormais para Favelas e Comunidades Urbanas.

Lagakos, D., A. M. Mobarak, and M. E. Waugh (2023). The welfare effects of encouraging rural-urban migration. *Econometrica 91*(3), 803–837.

Lessing, B. (2020). Conceptualizing criminal governance. *Perspectives on politics 19*(3), 854–873.

Michaels, G., D. Nigmatulina, F. Rauch, T. Regan, N. Baruah, and A. Dahlstrand (2021a). Planning ahead for better neighborhoods: Long-run evidence from tanzania. *Journal of Political Economy 129*(7), 2112–2156.

Michaels, G., D. Nigmatulina, F. Rauch, T. Regan, N. Baruah, and A. Dahlstrand (2021b). Planning ahead for better neighborhoods: Long-run evidence from tanzania. *Journal of Political Economy 129*(7), 2112–2156.

Monte, F., S. J. Redding, and E. Rossi-Hansberg (2018). Commuting, migration, and local employment elasticities. *American Economic Review 108*(12), 3855–3890.

Monteiro, J. and R. Rocha (2017). Drug battles and school achievement: evidence from rio de janeiro's favelas. *Review of Economics and Statistics 99*(2), 213–228.

Soto, H. et al. (2000). *The mystery of capital: why capitalism triumphs in the West and fails everywhere else*, Volume 63. Basic Books/New York.

Tsivanidis, N. (2023). Evaluating the impact of urban transit infrastructure: Evidence from bogotá's transmilenio. *Conditionally accepted, American Economic Review*.

Ulyssea, G. (2010). Regulation of entry, labor market institutions and the informal sector. *Journal of Development Economics 91*, 87–99.

Ulyssea, G. (2020). Informality: Causes and consequences for development. *Annual Review of Economics 12*(1), 525–546.

Ulyssea, G., M. Bobba, L. Gadenne, and M. Harari (2025, April). Informality. *VoxDevLit 6*(2).

UN (2024). The sustainable development goals report 2024. Technical report.

UN-Habitat (2003). The challenge of slums: global report on human settlements. *Management of Environmental Quality: An International Journal 15*(3).

Uribe, A., B. Lessing, N. Schouela, and E. Stecher (2025). Criminal governance in latin america: Prevalence and correlates. *Perspectives on Politics*, 1–19.

Zarate, R. D. (2024). Spatial misallocation, informality, and transit improvements: Evidence from mexico city. *Revise & Resubmit, American Economic Review*.