

On the Determinants of Slum Formation*

Tiago Cavalcanti Daniel Da Mata Marcelo Santos

Version: July 4, 2018.

Abstract

We construct a simple model of a city with heterogeneous agents and housing choice to explain the determinants of slums, home to about one-third of the urban population in developing countries. The model supports the main empirical evidence regarding slum formation and is able to quantitatively assess the role of each determinant of slum growth. We show that urban poverty, inequality, and rural-urban migration explain much of the variation in slum growth in Brazil from 1980 to 2000. Ex-ante evaluation of the impacts of policy interventions shows that removing barriers to formalisation has a strong impact on slum reduction.

Keywords: Slums; Growth; Rural-Urban Migration; Land Use Regulations.

JEL Classification: O18, O15, R52

*We thank Morten Ravn and two anonymous referees for their valuable comments and constructive suggestions on the paper. We also thank Toke Aidt, Klenio Barbosa, Thomas Crossley, Daniel Sturm, Leo Feler, Jane Fruehwirth, Petra Geraats, Chryssi Giannitsarou, Vernon Henderson, Pramila Krishnan, Alice Kuegler, Danilo Iglori, Povilas Lastauskas, Hamish Low, Diana Motta, Cezar Santos, Flavio Toxvaerd and Anne Villamil and the participants of several seminars for their useful comments. All remaining errors are our own. We thank the Keynes Fund, University of Cambridge, for financial support. Corresponding Author: Daniel Da Mata. Mailing Address: Rua Itapeva, 474 – 12 andar, Bela Vista, CEP 01332-000, Sao Paulo, Brazil. E-mail: daniel.damata@fgv.br

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as

doi: 10.1111/eoj.12626

This article is protected by copyright. All rights reserved.

1 Introduction

Slums represent a large proportion of housing markets in developing countries. Although the definition of slums varies depending on the country, slums are always associated with some sort of deprivation such as insecure land tenure, low standards of urban services, and non-durable housing structures.¹ Slums generally occupy marginal unauthorised land, mostly in environmentally vulnerable locations such as around garbage disposal sites, riverbanks, and slopes susceptible to landslides. Slum settlements usually have limited access to clean water, proper sewage treatment, or adequate heating, and therefore are associated with worse quality of life. Even though slums are subject to environmental hazards, constant raids and patrols, and social stigma, hundreds of millions of people worldwide migrate to cities and establish residence in slum areas. According to UN-Habitat (2016), in 2014 30% of the urban population in developing countries lived in slums. This corresponds to about 881 million people living in slums. Despite the large number of slum dwellers around the world, not much is known about the economic incentives associated with the decision to live in informal settlements, or about which public policies are most effective in altering slum formation (Brueckner and Selod (2009)).

In this paper, we address this gap and study the determinants of slums by investigating the economic incentives behind squatting. We build a general equilibrium model of a city with heterogeneous agents and housing choices. The model is consistent with the main empirical evidence on slum formation and can quantitatively assess the role of each determinant of slum growth. The model's structure is based on microdata from the Caju slum in Brazil.² The model is calibrated and estimated to be consistent with several statistics related to the Brazilian urbanisation process occurring from 1980 to 2000 (with a focus on the city of São Paulo). We use the Brazilian case to motivate the analysis and estimate the model, but the theoretical framework can be easily applied to other developing

¹We use informal housing, slums, and squatter settlements interchangeably. By and large, slums are characterised by overcrowding, lacking access to basic urban services, and an illegal occupation of land (UN-Habitat (2016)). Squatting involves illegally dwelling on land that one does not own (Jimenez (1985)). Informal housing can be seen as a general concept relating to conditions where there is insecure land tenure, non-compliance with building regulations, and/or inappropriate infrastructure.

²Caju slum data were collected for a land titling project and are detailed in online Appendix A.

countries. We perform several counterfactual event and policy simulations to disentangle the impacts of specific factors on informal housing growth.

In our model, households are heterogeneous in their labour productivity, and they have to decide between two housing tenure modes (formal or informal housing). In making decisions on tenure modes, households face the following trade-off: if they choose to live in an informal settlement, they avoid paying property taxes and complying with building regulations. However, informal housing is insecure (due to, e.g., the exogenous risks of eviction and demolition by the government or property theft by other residents); therefore, they incur (i) a utility reduction per unit of housing space and (ii) protection costs. We model protection costs as forgone labour income: agents must spend part of their time physically present to protect their informal plots.³

The basic economic mechanism generates two income thresholds, that separate formal from informal housing agents. The first cut-off comes from the opportunity cost of protecting the informal plot. As income grows, protection costs in the form of forgone labour income increase, and therefore households are better off living in the formal housing sector. Thus, more affluent individuals tend to live in formal housing units (e.g., Friedman, Jimenez and Mayo (1988)). The second cut-off is generated by zoning constraints that interfere with households' decisions and indicates how housing regulations may induce slum formation. Households that choose to live in formal housing units must comply with several building constraints, such as minimal lot sizes, height restrictions, and green area ratios, among other building regulations. Households unable to meet these requirements are bound to live in informal settlements. The model thus reveals two reasons why for poor households the (only) feasible option is to live in slums: to avoid binding building regulations and on costs related to formal housing.

Our model is able to account for a large proportion of the variation in slum growth occurring over the past several decades in Brazil. Interestingly, changes in the income distribution (mean and dispersion) and rural-urban migration explain much of the variation

³Apart from making up for lack of police protection, the time costs of being in a slum can include the time required to recover from health problems related to poor urban infrastructure, the time cost of getting potable water, and paying fees to a community leader.

in slum growth occurring between 1980 and 2000. Our counterfactual exercises focus on two types of policies: policies that weaken barriers to formalisation and slum upgrading interventions. Our simulations show that decreasing barriers to formalisation (e.g., adopting housing-friendly land use regulations and decreasing property taxes) has a strong impact on slum reduction. In contrast, slum upgrading interventions can have unintended effects, as incomplete upgrading programs (e.g., providing basic infrastructure but maintaining non-compliance with existing building regulations) can make informal housing more attractive and induce more people to live informally. Additionally, we perform a welfare exercise to understand the distributional impacts of the simulated policies. Reducing property taxes has a relatively large impact on welfare according to the policy simulations. The main lessons drawn from the simulations are that (i) one must differentiate between types of policies and (ii) the average welfare effects are mainly driven by the subset of the population that is directly affected by each policy intervention, even though a large proportion of the population may be affected by the policy due to the presence of pecuniary externalities.

The remainder of this section discusses the related literature and highlights our contributions. The motivated evidence is presented in Section 1. Section 2 presents our model. In Section 3 we discuss steps taken to calibrate the model and to estimate its parameters. Section 4 presents our quantitative analysis on the determinants of slum formation and our ex-ante policy simulations. Section 5 contains concluding remarks.

Literature.

Our article follows a literature that quantitatively investigates how different policies affect urban phenomena. See, for instance, Chatterjee and Carlino (2001), Baum-Snow (2007), Kopecky and Suen (2010), Michaels, Rauch and Redding (2012), and Brinkman, Coen-Pirani and Sieg (2015). Our paper differs from these papers in that none of them have investigated the issue of informal housing.⁴

Our paper is also related to the literature on urban squatting and informal housing.

⁴There is also a literature that investigates the determinants of informal production in developing countries (e.g., Amaral and Quintin (2006) and Antunes and Cavalcanti (2007)). In this literature, the entrepreneurial choice of households depends on their entrepreneurial income while in our model the type of housing choice of households depends on their labour productivity.

Field (2007) provides evidence of the importance of forgone labour income when agents lack property titles. Brueckner and Selod (2009) and Brueckner (2013) present models of community organisers and informal housing. Cai, Selod and Steinbuks (2018) show that the costly provision of property rights can lead to slum persistence. The main differences between our paper and those on urban squatting are as follows: (i) we set up a general equilibrium model of a city with heterogeneous agents, while the literature typically employs a partial equilibrium analysis, or a representative agent setting; and (ii) we not only derive the qualitative properties of the model, but we also study it quantitatively by estimating parameters of the model based on evidence from Brazil and implementing an ex-ante policy evaluation.

Finally, our focus on policy evaluation links our paper to existing papers on the effects of slum upgrading. This literature has focused more heavily on the ex-post evaluation of policies. Examples include studies on how land titles and infrastructure upgrades influence slum outcomes in Latin American countries. See for instance Lanjouw and Levy (2002), Field (2005), Field (2007), Di Tella, Galiani and Schargrodsky (2007), Galiani and Schargrodsky (2010) and Galiani, Gertler, Undurraga, Cooper, Martinez and Ross (2017) among others. We aim to contribute by performing an ex-ante evaluation of key policies on slums and understanding the main mechanisms underlying the results.

2 Data Facts and Reduced-Form Evidence

In this section, we document some facts from Brazil that motivated this study and provide empirical support for some of our modelling strategies. In Brazil, there is no official definition of a “slum”, but a working definition called of “subnormal agglomeration” is obtained from the population censuses. According to the census framework, a subnormal agglomeration satisfies three conditions: (i) it includes a group of at least 50 housing units (ii) that occupies land illegally and (iii) is urbanised in a disordered fashion and/or that lacks basic public services such as sewage or electricity services. Therefore, there is a connection between the definition of “subnormal agglomeration” and the notion of a

“slum”. Data sources and definitions are shown in the companion online Appendix A.

2.1 Data Facts

Fact 1: Brazil has undergone rapid urbanisation in recent decades. The total population jumped from approximately 50 million in 1950 to over 190 million in 2010, and the share of its urban population increased from 36% to 84% in the same period (see Table A2 in online Appendix A). This surge in urban population has been accompanied by stagnant real incomes and persistent income inequality. In the mid-2000s, real wages recovered to the level of the 1980s after a sharp drop in the early 1990s (see Figure A1 in online Appendix A). Brazil presents very high and persistent levels of income inequality, although they have recently declined.

Fact 2: The last few decades have also been associated with the adoption of stricter land use regulations by Brazilian municipalities.⁵ In 1979, Federal Law 6,766 stipulated that developers in Brazil must follow federally mandated construction parameters on basic infrastructure and land parcelling. One such country-wide parameter applied a minimum lot size of 125 m². The federal law also allowed municipalities to alter such parameters by creating local land use laws based on locally suitable parameters. Several municipalities have enacted local land use laws, but instead of lowering the federal requirements, the majority established stricter construction parameters that made formal land less affordable. From 1979 to 1999, of the 1,482 municipalities that had established their own land use laws, 1,183 set larger minimum lot sizes.⁶

Fact 3: There has been a rise in the number of slums and slum dwellers in Brazilian cities. Population census data (1980-2010) show that slums are steadily growing in Brazil (see Table 1). Even though census data are capable of capturing slum growth trends, they tend to underestimate the number of slum units and slum dwellers present. The Underestimation of slum size is a common problem facing developing countries (Marx,

⁵A municipality is considered an (autonomous) local administrative unit in Brazil. They are roughly equivalent to counties in the United States.

⁶In fact, lowering the national subdivision standards is considered unconstitutional in Brazil — except for specific urbanisation projects to build public housing.

Stoker and Suri (2013)).

Table 1: Population, Housing and Slums in Brazil: 1980-2010

123 urban agglomerations	1980	1991	2000	2010
Housing units(a)	14,012,484	20,564,931	27,126,584	34,188,992
Slum units(b)	476,292	943,667	1,488,779	2,732,576
Population(c)	62,390,783	80,885,091	96,951,317	110,287,840
Slum dwellers(d)	2,224,164	4,084,051	5,775,890	9,516,899
% Slums(b/a)	3.40%	4.59%	5.49%	7.99%
% Slum dwellers (d/c)	3.56%	5.05%	5.96%	8.63%

Notes. Data from Da Mata, Lall and Wang (2008) and tabulations from the 2010 Population Census in Brazil. Slum data from the 2010 Census is not directly comparable with that of other censuses (IBGE (2011)). The data are for 123 urban agglomerations in Brazil.

2.2 Reduced-Form Evidence: Correlations

The data suggest that changes in income, population, and regulation may have important implications for the evolution of slums in Brazilian cities. Now, we present reduced-form evidence that aims to relate these facts by estimating the following specification:

$$y_{it} = \beta' X_{it} + \epsilon_{it},$$

where y_{it} is the share of slum dwellers in city i in year t , X_{it} are (potential) city-level determinants of slum growth, β is a vector of parameters to be estimated, and ϵ_{it} represents the unobservable determinants of slum growth.

Table 2 shows correlations regarding the potential determinants of slum formation at the city level: *per capita* income, income inequality (Gini coefficient), urban regulation, and population density. We use panel data for 123 urban agglomerations in Brazil for 1980-2000.⁷ There is a positive correlation between slum formation and three potential determinants, suggesting that higher levels of income inequality, stricter regulation, and larger urban populations are associated with the presence of more slums. After controlling for other factors, the mean *per capita* income is either positively correlated with slum growth (see columns (iv) and (vii)) or is not statistically significant (see column (v)). One

⁷We do not use data for 2010 in our panel analysis because slum data of the 2010 Census are not directly comparable with those of other censuses (IBGE (2011)). See Da Mata, Deichmann, Henderson, Lall and Wang (2007) for a definition of the 123 urban agglomerations. See the companion online Appendix A for more details.

possible explanation is that places with higher *per capita* incomes have higher housing prices, which can act as a barrier to formal housing market entry. To check the robustness of our results, we test each potential determinant separately and their combinations (see columns (i) to (v) of Table 2). We also employ a different dependent variable - the (log) number of slum dwellers in the urban agglomeration (see columns (vi) and (vii) of Table 2) - and the results are similar. We perform additional robustness exercises. The correlations remain mostly unchanged with random-effects but are less robust with fixed-effects. A more complete discussion is provided in online Appendix B.

Table 2: Reduced-Form Evidence: Pooled OLS

Variables	% of slum dwellers					ln # of slum dwellers	
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
(ln) <i>per capita</i> income	0.0105** (0.0045)			0.0156** (0.0064)	0.0092 (0.0060)	3.507*** (0.841)	2.337*** (0.851)
Income inequality (Gini index)		0.172*** (0.0482)		0.235*** (0.0526)	0.257*** (0.0507)	35.99*** (6.991)	40.02*** (6.435)
Urban regulation			0.0119*** (0.0029)	0.0117*** (0.0041)	0.0102** (0.0039)	1.709** (0.704)	1.444** (0.702)
(ln) Population density					0.00636*** (0.0018)		1.161*** (0.223)
Constant	-0.0430* (0.025)	-0.0775*** (0.026)	0.0056*** (0.002)	-0.210*** (0.054)	-0.215** (0.051)	-36.69*** (6.911)	-37.63*** (6.690)
Observations	369	246	369	246	246	246	246
R-squared	0.013	0.038	0.017	0.077	0.127	0.122	0.216

Notes. Robust standard errors in parentheses. All the regressions are for 123 urban agglomerations in Brazil. The pooled OLS is for 1980, 1991, and 2000. Therefore, there are 369 observations in each regression. The dependent variables are the share of slum dwellers in each urban agglomeration (columns (i) to (v)) and the log number of slum dwellers (columns (vi) and (vii)). Each urban agglomeration has more than one municipality. The explanatory variable “urban regulation” is a dummy variable that takes value 1 if the majority of people in the urban agglomeration live in municipalities with their own land use law. The Gini index is measured only in two years (1991 and 2000), explaining the 246 observations in regressions (ii), (iv), (v), (vi) and (vii).

*** p<0.01, ** p<0.05, * p<0.1

The coefficients of the reduced-form results give us the directions of the correlations, but the exact mechanisms explaining slum growth remain unclear. In the next section, we present a model for studying these mechanisms.

3 The Model

We present an equilibrium model of a city to explain the determinants of informal housing formation. The model includes households, with distinct income levels due to heteroge-

neous labour productivity, that choose to live in either formal or informal housing. The trade-off between the two tenure modes is as follows. Households living in informal housing settlements do not need to comply with building regulations, and can avoid the taxes and fees associated with their housing. Informal housing, however, lacks access to public infrastructure and is insecure; thus, informal dwellers incur time costs to protect their plots. We use microdata on the Caju slum in Brazil to develop the model's structure.

In this section, we keep the model as simple as possible so the analytical results can be derived. We perform a quantitative analysis of this baseline model in Sections 3 and 4. In online Appendix D, we relax some assumptions of the baseline model and compare the results. We show that the extended model gives qualitatively similar results.

3.1 The Environment

Households. There is a continuum of households of measure P , each with labour productivity $\lambda \in (0, \infty)$, that follows a continuous cumulative distribution $\Upsilon(\lambda)$ where $\Upsilon(0) = 0$ and $\lim_{\lambda \rightarrow \infty} \Upsilon(\lambda) = 1$. These households have utility over non-housing consumption (c) and housing services (s_j), where j equals F if formal housing is used and I if informal housing is used. The utility function is specified as

$$U(c, s_j) = c^{1-\alpha} s_j^\alpha, \quad \alpha \in (0, 1).$$

Housing service flow (s_j) entails the housing space (h) and public goods surrounding a housing unit (g) such that:

$$s_j = \begin{cases} gh & \text{if formal housing (j=F),} \\ \theta gh & \text{if informal housing (j=I),} \end{cases}$$

where $\theta \in (0, 1)$. Expression θgh means that informal housing agents do not reap all benefits from public goods provisions within city boundaries. Parameter θ may be seen as a congestion cost incurred by living in high-density areas. The Caju slum data support the

low provision of public services and high congestion costs incurred in informal settlements.

Households' labour incomes are given by $(1 - \tau)w\lambda$, where w is the average wage of the given city, τ is labour income tax, and λ is the labour productivity of a household. A household allocates labour income to non-housing consumption (c), housing services (s), and a property tax (η). Only households living in formal housing pay a property tax η — a cost associated with being located in the formal sector. One can argue that slum dwellers also avoid paying for public utilities such as electricity services. Therefore, parameter η can also consider the payment of public utilities by formal housing residents, whereas informal housing residents do not pay for such public services. Approximately 94% of the households in Caju do not pay property taxes, but 92% of the residents in Caju pay electricity bills.

A household has one unit of productive time. If a household decides to live in the informal sector, it incurs protection costs (Ψ). Informal housing lacks property rights and is associated with insecurity due to the exogenous risks of eviction and demolition from the government and even a risk associated with property theft by other households. Therefore, if an agent chooses to dwell illegally, she will spend $1 - \Psi$ of her time working; i.e., informal housing agents allocate part of their time to protecting their informal plots. The implicit assumption is that if households living in informal housing do not spend time or resources protecting their land, then they will be evicted. This implies that spending such time and resources is incentive compatible. Protection costs are given by $\Psi(\lambda, h) = \psi(1 - \tau)w\lambda h$. This implies that the opportunity cost associated with the protection of a plot increases with income and housing space.⁸ Parameter $\psi > 0$ reflects both penalty costs and the probability of being penalised by the government for living in an informal settlement.

A formal housing unit must occupy a space that must comply with a specific building-related requirement: the minimum lot size (MLS). This is a catch-up parameter that summarises several building regulations related to formal housing, such as construction standards, height restrictions and drainage systems. In the model, MLS is represented by

⁸Instead of forgone labour, the opportunity cost of informal housing can be modelled as defensive expenditures paid to a community leader, for instance. This assumption would generate similar results.

the parameter $\underline{h} \geq 0$, which is exogenously set by the local government.

In sum, if choosing to live in an informal settlement, households avoid paying property taxes (η) and complying with minimum lot size regulations (\underline{h}), but informal housing is insecure. Such a household incurs a utility discount per unit of housing space (θ) and an opportunity cost related to informal housing (ψ). As a result, the informal housing sector in our model is related to the general notion of a slum, where there is insecure land tenure (represented by the opportunity cost ψ), poor infrastructure provisions (represented by θ), and non-compliance with building regulations (represented by \underline{h}).

Consumption goods. The consumption good is the numeraire. Its production takes place in a competitive environment in which all firms have access to the production function $Q(N, K) = BN^v K^{1-v}$, where $v \in (0, 1)$ is the labour share in production, N denotes units of labour efficiency, K represents units of capital and B is the total factor productivity parameter. Labour is supplied by households. The rental price of capital r is determined outside the city. The optimisation problem of the representative firm operating in this city gives the labour demand of the city:

$$w = vB \left[\frac{(1-v)B}{r} \right]^{\frac{1-v}{v}}. \quad (1)$$

Housing and the city. Land L is available in fixed supply. The production function of housing space is given by $H(L, K) = L^\gamma K^{1-\gamma}$ where $\gamma \in (0, 1)$. Let the rental price of land be given by p_L . Land belongs to a group of absentee landlords who spend land rents outside the city. Developers solve the following optimisation problem:

$$\Pi = \max_{L, K} \{pL^\gamma K^{1-\gamma} - p_L L - rK\},$$

where p is the housing price. From the first-order conditions for this problem, the following applies:

$$p_L = \gamma \left(\frac{1-\gamma}{r} \right)^{\frac{1-\gamma}{\gamma}} p^{\frac{1}{\gamma}}. \quad (2)$$

Housing supply is obtained by manipulating the first-order conditions of this problem:

$$H^s = Lp^{\frac{1-\gamma}{\gamma}} \left[\frac{1-\gamma}{r} \right]^{\frac{1-\gamma}{\gamma}}, \quad (3)$$

which is a constant price-elasticity ($\frac{1-\gamma}{\gamma}$) housing supply function.

Government and land law enforcement. Public goods g are financed through a property tax (η) paid by formal housing residents and a labour income tax (τ) paid by all workers. Define E^F (see in short the formal definition) as the measure for households residing in formal housing units. The government runs a balanced budget such that

$$g = \int_{E^F} \eta R_H h_F(\lambda) d\Upsilon(\lambda) + \int_0^\infty \tau w \lambda d\Upsilon(\lambda), \quad (4)$$

where R_H is the housing rent and $h_F(\lambda)$ is the formal housing size chosen by individual λ .

3.2 Decision Problems and Properties of the Model

Formal housing problem. Taking w (average city wage) and R_H (housing rent) as given, households that have chosen to live in formal housing solve the following problem:

$$V^F(\lambda) = \max_{c \geq 0, h} [c^{1-\alpha} (gh)^\alpha], \quad (5)$$

subject to

$$c + [R_H(1 + \eta)]h \leq w\lambda(1 - \tau), \quad (6)$$

$$h \geq \underline{h} \geq 0. \quad (7)$$

The budget constraint (6) states that expenditures on consumption and housing cannot be greater than disposable income. Equation (7) shows that there is a minimum formal housing space, that may be greater than zero. When equation (7) is not binding, optimal consumption and housing choices for formal households are determined by

$$c_F(\lambda) = (1 - \alpha)w\lambda(1 - \tau), \quad (8)$$

$$h_F(\lambda) = \alpha \frac{w\lambda(1-\tau)}{R_H(1+\eta)}. \quad (9)$$

Substituting these two equations into the utility function yields the following value function:

$$V^F(\lambda) = V_{nb}^F(\lambda) = \alpha^\alpha (1-\alpha)^{1-\alpha} w(1-\tau)\lambda [R_H(1+\eta)]^{-\alpha} g^\alpha. \quad (10)$$

When equation (7) is binding, we find that

$$c_F(\lambda) = w\lambda(1-\tau) - R_H(1+\eta)\underline{h}, \quad (11)$$

$$h_F(\lambda) = \underline{h}. \quad (12)$$

In this case, the indirect utility function of a household is given by

$$V^F(\lambda) = V_b^F(\lambda) = [w(1-\tau)\lambda - R_H(1+\eta)\underline{h}]^{1-\alpha} \underline{h}^\alpha g^\alpha. \quad (13)$$

Constraint (7) binds only to sufficiently small values of λ . Moreover, there is a minimum λ , that is equal to $\underline{\lambda} = \frac{R_H(1+\eta)\underline{h}}{w(1-\tau)} \geq 0$ such that for any $\lambda < \underline{\lambda}$ the problem of a household residing in formal housing is not well defined. Then, for any $\lambda > \underline{\lambda}$ we have that

$$h_F(\lambda) = \max \left[\alpha \frac{w\lambda(1-\tau)}{R_H(1+\eta)}, \underline{h} \right]. \quad (14)$$

The problem of a household residing in formal housing is formally characterised in Lemma C.1 in online Appendix C. The black solid line shown in Figure 1 depicts $V^F(\lambda)$.⁹

Informal housing problem. Households choosing to live in informal housing solve the following:

$$V^I(\lambda) = \max_{c,h \geq 0} [c^{1-\alpha}(\theta gh)^\alpha], \quad (15)$$

subject to

$$c + R_H h \leq w\lambda(1-\tau) - \Psi(\lambda, h), \quad \text{with } \Psi(\lambda, h) = \psi(1-\tau)w\lambda h. \quad (16)$$

⁹ $V^F(\lambda)$ is continuous and strictly increasing in λ . $V_{nb}^F(\lambda)$ and $V_b^F(\lambda)$ are also displayed in this figure. $V_{nb}^F(\lambda)$ is represented by the dashed black line, and $V_b^F(\lambda)$ is represented by the dotted black line.

This problem leads to the following consumption and housing demand equations:

$$c_I(\lambda) = (1 - \alpha)w\lambda(1 - \tau), \quad (17)$$

$$h_I(\lambda) = \alpha \frac{\lambda w(1 - \tau)}{R_H + \psi w\lambda(1 - \tau)}. \quad (18)$$

The indirect utility function for households living in informal housing is given by

$$V^I(\lambda) = \alpha^\alpha (1 - \alpha)^{1 - \alpha} w(1 - \tau)\lambda [R_H + \psi w(1 - \tau)\lambda]^{-\alpha} (g\theta)^\alpha. \quad (19)$$

$V^I(\lambda)$ is depicted by the solid grey line shown in Figure 1 and formally described in Lemma C.2 in online Appendix C. Clearly, when $\lambda \in (\underline{\lambda}, \lambda^{MLS})$, formal housing has an income elasticity of zero; for $\lambda > \lambda^{MLS}$, then formal housing has an income elasticity of 1. Informal housing has an income elasticity of less than 1 (see Proposition C.3 in online Appendix C). These results are consistent with the literature on the lack of incentives available to invest in informal housing units (Kapoor and le Blanc (2008)).

Housing choice. The housing type decision can be described by the indicator function $\Omega(\lambda) \in \{0, 1\}$, where $\Omega(\lambda) = 1$ whenever $V^F(\lambda) > V^I(\lambda)$ and zero otherwise. Define the measure of households who choose to live in a formal housing by

$$E^F = \{\lambda \in (0, \infty) : V^F(\lambda) > V^I(\lambda)\}, \quad (20)$$

and let $(E^F)^c$ denote the complementary set of E^F . Clearly, if $\lambda \in (E^F)^c$, then households are residing in informal housing settlements.

ASSUMPTION 1. *Let $(1 + \eta)\theta > 1$.*

Assumption 1 guarantees that very poor households prefer to live in informal housing than in formal housing when $\underline{h} = 0$. If Assumption 1 does not hold, slums would exist only due the presence of binding regulations in the formal housing market. The following lemma characterises the housing choice for a given labour ability and prices:

LEMMA 1 *Let Assumption 1 be satisfied. Then, there exists a unique $\bar{\lambda}$, such that any*

$\lambda \in (0, \bar{\lambda})$ implies $\lambda \in (E^F)^c$; any $\lambda > \bar{\lambda}$ implies $\lambda \in E^F$. Moreover,

i. $\frac{\partial \bar{\lambda}}{\partial \theta} > 0$; $\frac{\partial \bar{\lambda}}{\partial \eta} > 0$; $\frac{\partial \bar{\lambda}}{\partial \psi} < 0$; $\frac{\partial \bar{\lambda}}{\partial h} \geq 0$;

ii. $\frac{\partial \bar{\lambda}}{\partial w} < 0$; and $\frac{\partial \bar{\lambda}}{\partial R_H} > 0$.

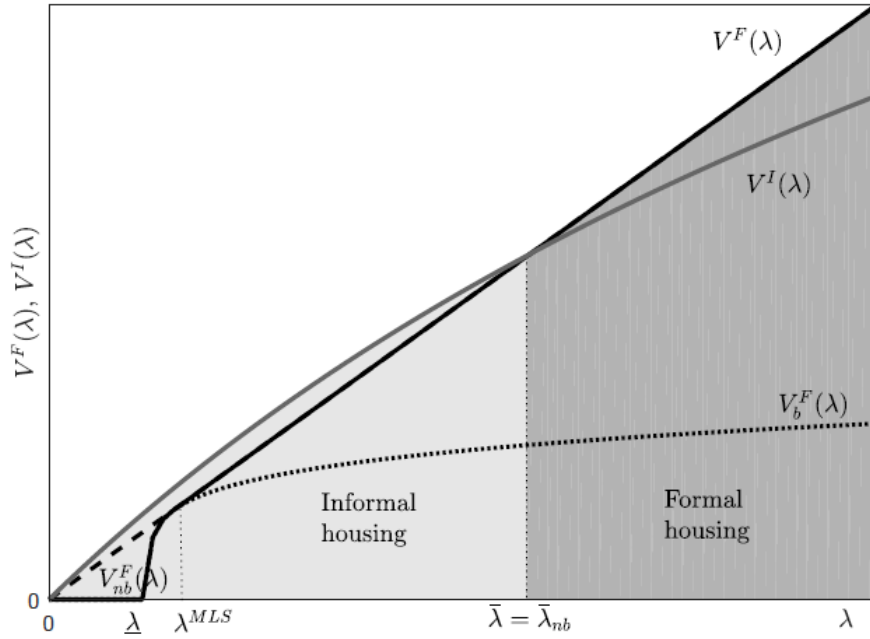
Proof. Find $\bar{\lambda}$ by using $V^F(\lambda) - V^I(\lambda) = 0$. Then, use the Implicit Function Theorem to investigate how $\bar{\lambda}$ changes with parameters and prices. See the companion online Appendix C for details. ■

According to Assumption 1, $V^F(\lambda) - V^I(\lambda) = 0$ defines a unique value of $\bar{\lambda} > 0$ for labour productivity such that households are indifferent to living formally or informally. For any $\lambda \in (0, \bar{\lambda})$ households choose to live informally, and for any $\lambda > \bar{\lambda}$ households prefer to live in formal housing. Two cases must be considered. First, the productivity level λ^{MLS} associated with the minimum lot size constraint is lower than $\bar{\lambda}$. In turn, housing regulations do not interfere with the housing decisions of households. This case is depicted in Figure 1(a) and

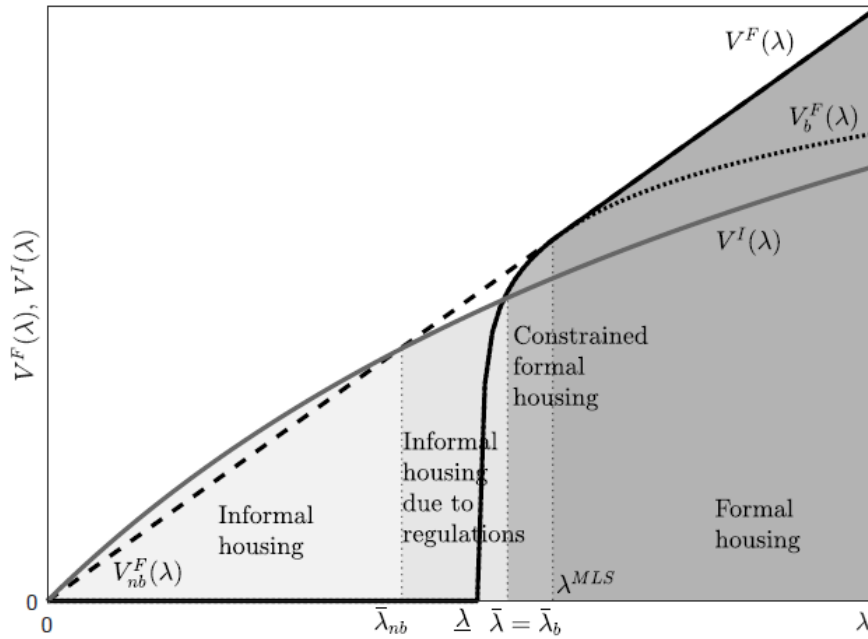
$$\bar{\lambda} = \bar{\lambda}_{nb} = \frac{R_H}{\psi w(1 - \tau)}[\theta(1 + \eta) - 1]. \quad (21)$$

The region of formal housing choice is presented by the dark grey area, and informal housing is characterised by the light grey area. The share of informal housing increases with housing rent levels (R_H), property taxes (η) and with low costs of informality (high θ).

Fig. 1: Housing Choice



(a) Non-binding housing regulations. λ corresponds to the lowest productivity value in which $V_b^F(\lambda)$ is defined; λ^{MLS} is the productivity value such that for any $\lambda \leq \lambda^{MLS}$ the minimum size regulation is binding in the formal sector. Therefore, $V^F(\lambda) = V_b^F(\lambda)$ for $\lambda \in (\lambda, \lambda^{MLS})$ and $V^F(\lambda) = V_{nb}^F(\lambda)$ for $\lambda > \lambda^{MLS}$. The cut-off $\bar{\lambda}_{nb}$ ($\bar{\lambda}$) is the productivity value such that $V_{nb}^F(\bar{\lambda}) > V^I(\bar{\lambda})$ ($V^F(\bar{\lambda}) > V^I(\bar{\lambda})$) for any $\lambda > \bar{\lambda}_{nb}$.



(b) Binding housing regulations. λ corresponds to the lowest productivity value in which $V_b^F(\lambda)$ is defined; λ^{MLS} is the productivity value such that for any $\lambda \leq \lambda^{MLS}$ the minimum size regulation is binding in the formal sector. Therefore, $V^F(\lambda) = V_b^F(\lambda)$ for $\lambda \in (\lambda, \lambda^{MLS})$; $\bar{\lambda}_{nb}$ is the productivity value such that $V_{nb}^F(\bar{\lambda}_{nb}) > V^I(\bar{\lambda}_{nb})$; $\bar{\lambda}_b$ ($\bar{\lambda}$) is the productivity value such that $V_b^F(\bar{\lambda}_b) > V^I(\bar{\lambda}_b)$ for any $\lambda > \bar{\lambda}_b$.

When $\lambda^{MLS} > \bar{\lambda}$, regulation binds for tenure type decisions. Some agents choose to live informally (because they cannot comply with the minimum lot size), even though they would choose to live in the formal housing sector if such regulations were not in place. This case is depicted in Figure 1(b) for households with $\lambda \in [\bar{\lambda}_{nb}, \bar{\lambda})$, where $\bar{\lambda}_{nb}$ denotes the ability level such that households would be indifferent to choosing between formal and informal housing if there were no *MLS* regulations. Households with ability $\lambda \in [\bar{\lambda}, \lambda^{MLS})$ live in formal housing but the size of their housing choices is constrained by *MLS* regulations. The application of more restricted housing regulations (higher value of \underline{h}) implies the presence of a large share of informal housing.

When Assumption 1 is not satisfied, $V_{nb}^F(\lambda) > V^I(\lambda)$ for any value of $\lambda > 0$ and informal housing exists only if there are *MLS* regulations. This will be similar to the case depicted in Figure 1(b), except that $V^I(\lambda)$ will be uniformly lower than $V_{nb}^F(\lambda)$.

3.3 Equilibrium

To complete the analysis of the model, it remains to determine five (w, r, p_L, R_H, p) prices of this economy. The rental price of capital is determined outside the city boundaries. Therefore, given r and Equation (1), the wage rate w is also determined. Equilibrium in the asset market with no uncertainty means that the housing price equals the discounted value of housing rents: $p = R_H/r$, where R_H represents the housing rent. Equation (2) implies that given r and once we determine p , the land price p_L can be determined. Therefore, we just need one market clearing condition to define the housing price p , and, thus, the land price, p_L , and housing rent, R_H .

Housing market equilibrium. The total housing demand in the informal sector is

$$H^I(R_H) = P \int_0^{\bar{\lambda}} h_I(\lambda) d\Upsilon(\lambda) = P \int_0^{\bar{\lambda}} \alpha \frac{w\lambda(1-\tau)}{R_H + \psi w\lambda(1-\tau)} d\Upsilon(\lambda).$$

In the formal sector, the total housing demand is defined as

$$H^F(R_H) = P \int_{\bar{\lambda}}^{\infty} h_F(\lambda) d\Upsilon(\lambda) = P \int_{\bar{\lambda}}^{\infty} \max \left[\alpha \frac{w\lambda(1-\tau)}{R_H(1+\eta)}, \underline{h} \right] d\Upsilon(\lambda).$$

Recall that housing supply is given by Equation (3) and $p = R_H/r$. The housing market clearing condition can be defined as

$$P \int_0^{\bar{\lambda}} \alpha \frac{w\lambda(1-\tau)}{R_H + \psi w\lambda(1-\tau)} d\Upsilon(\lambda) + P \int_{\bar{\lambda}}^{\infty} \max \left[\alpha \frac{w\lambda(1-\tau)}{R_H(1+\eta)}, \underline{h} \right] d\Upsilon(\lambda) = LR_H^{\frac{1-\gamma}{\gamma}} \left[\frac{1-\gamma}{r^2} \right]^{\frac{1-\gamma}{\gamma}}, \quad (22)$$

where the left-hand side (LHS) corresponds to the total housing demand and the right-hand side (RHS) denotes the supply of housing in the city.

PROPOSITION 1 (*Existence and Uniqueness*) *There exists a rental housing price level $0 < R_H^* < \infty$ that clears the housing market. Moreover,*

- i. When $\lambda^{MLS} < \bar{\lambda}_{nb}$, then a sufficient condition for uniqueness is that $\theta \rightarrow 1$.*
- ii. When $\lambda^{MLS} > \bar{\lambda}_{nb}$, then a sufficient condition for uniqueness is that $\underline{h} \geq \frac{\eta}{(1+\eta)} \frac{\alpha}{\psi}$.*

Proof. Excess housing demand is continuous in R_H . In addition, it is strictly positive for small values of R_H , and it is strictly negative for large values of R_H . Therefore, an equilibrium exists. The derivative of the excess housing demand with respect to R_H can be negative or positive, and it is strictly negative when conditions stated in (i) and (ii) above are satisfied, respectively. See the companion online Appendix C for details. ■

Given the properties of this model, an equilibrium always exists. The conditions for uniqueness are sufficient. Even when these conditions are not satisfied, the equilibrium may be unique. In fact, in most of our numerical exercises we find a unique equilibrium price. Note that uniqueness does not mean that the “effective costs” of formal and informal housing units are equal: for each agent, the cost of a housing unit *differs* on her ability. Specifically, for a given R_H , an informal housing unit is “cheaper” for low-income agents given low opportunity costs required for this household to protect the plot.

We can now study the qualitative properties of the model, which are summarised in Proposition 2. For tractability, in this proposition, we assume the government keeps the same level of public goods. The government’s budget constraint is adjusted by changing the fraction of resources that may be wasted.¹⁰ In the numerical exercises, we also allow

¹⁰In this case, the government budget constraint is given by $g + \Delta = \int_{E^F} \eta R_H h_F(\lambda) d\Upsilon(\lambda) + \int_0^{\infty} \tau w \lambda d\Upsilon(\lambda)$ and Δ is adjusted.

changes in policies and institutions to affect the level of public good provision.

PROPOSITION 2 (*Comparative statics*) *Let the conditions for uniqueness be satisfied, then*

- i. $\frac{\partial R_H^*}{\partial P} > 0$; $\frac{\partial R_H^*}{\partial L} < 0$; $\frac{\partial R_H^*}{\partial r} > 0$; $\frac{\partial R_H^*}{\partial \theta} < 0$; $\frac{\partial R_H^*}{\partial w(1-\tau)} > 0$; $\frac{\partial R_H^*}{\partial \underline{h}} \geq 0$; $\frac{\partial R_H^*}{\partial \psi} \leq 0$; $\frac{\partial R_H^*}{\partial \eta} < 0$;
- ii. $\frac{\partial \bar{\lambda}}{\partial P} > 0$; $\frac{\partial \bar{\lambda}}{\partial L} < 0$; $\frac{\partial \bar{\lambda}}{\partial r} > 0$; $\frac{\partial \bar{\lambda}}{\partial \theta} > 0$; $\frac{\partial \bar{\lambda}}{\partial w(1-\tau)} \leq 0$; $\frac{\partial \bar{\lambda}}{\partial \underline{h}} \geq 0$; $\frac{\partial \bar{\lambda}}{\partial \psi} < 0$; $\frac{\partial \bar{\lambda}}{\partial \eta} > 0$;
- iii. *Let $\theta \rightarrow 1$ and let $\Upsilon^1(\lambda)$ and $\Upsilon^2(\lambda)$ be two continuous cumulative distributions with $\lambda \in (0, 1)$. If Υ^1 first-order stochastically dominates (FSD) Υ^2 or if Υ^1 second-order stochastically dominates (SSD) Υ^2 , then $R_H^{1,*} \geq R_H^{2,*}$ and $\bar{\lambda}^1 \geq \bar{\lambda}^2$.*

Proof. Use the housing excess demand function and condition $V^F(\bar{\lambda}) - V^I(\bar{\lambda}) = 0$ to define R_H^* and $\bar{\lambda}$. Then use the Implicit Function Theorem to show how R_H^* and $\bar{\lambda}$ change with each parameter of the model. See the companion online Appendix C for details. ■

The comparative statics help to understand the correlations found from the reduced-form evidence (Section 1). We find that the equilibrium rental housing price increases with population (P), interest rates (r), and disposable income ($w(1 - \tau)$). It decreases with the supply of land (L), when the utility cost for informality decreases (θ increases), and with property taxes (η). Such results are quite intuitive: larger populations and higher incomes, for instance, increase the demand for housing, with a positive effect on housing prices. We cannot identify how the rental housing price changes with protection costs (ψ) and minimum lot size regulations (\underline{h}) – it is unchanged when $\underline{h} > 0$ is not binding. For instance, when protection costs (ψ) decrease, more agents start to live informally, and there is an increase in housing size for informal households. The demand for formal housing, however, decreases, generating an ambiguous effect on the rental housing price.

The fraction of households living in informal housing increases for a larger population (P), lower supply of land (L), higher interest rates (r), lower utility cost of living informally (higher θ), higher property taxes (η), higher *MLS* regulations (if they are binding), and lower protection costs (ψ). We are not able to qualitatively identify how the share of households living in informal housing changes with disposable income. When wages increase, formal housing becomes more affordable and the cost of being a slum dweller

increases. This has a positive effect on the share of households living in formal housing (see item (ii) of Lemma 1). By contrast, the land price increases with higher housing demand, and this has a negative effect on the share of households living in formal housing. Therefore, the increase in land prices seems to explain the positive correlation between income and informality shown by the regressions of Table 2 in Section 1.¹¹

Finally, for the last item of Proposition 2, the distribution of abilities does not directly change the productivity threshold $\bar{\lambda}$ (see Lemma 1). Therefore, the ability distribution affects informal housing only through a general equilibrium effect. FSD and SSD imply a higher housing price with higher demand for housing spaces, which increases the threshold of the ability distribution. Note that we cannot state that informality rises because the mass of households in each ability interval also changes with different distributions. For instance, FSD means that a higher average productivity would increase the land price and $\bar{\lambda}$ would move to the right, thereby increasing housing informality. By contrast, the mass of agents at productivity $\lambda \geq \bar{\lambda}$ would also increase, thereby decreasing informal housing. However, a mean-preserving spread of the ability distribution (which implies SSD) would have a positive effect on informality in the housing market.

4 Estimation

The model formalises the mechanisms through which income, population size, and other variables affect housing type decisions. Now, we investigate the quantitative importance of these factors for slum formation. To accomplish this, we calibrate the model to match some moments of the city of São Paulo. We calibrate 14 parameters. In this section, we explain how their corresponding values were chosen.

External calibration. The following parameters were selected via external calibration: city-level productivity, B , labour share in the production of consumption goods, v , mean labour productivity, μ , variance in labour productivity, σ , weight of housing consumption, α , weight of land inputs, γ , property tax rates, η , income tax rates, τ , total urban

¹¹The model explains a case in which slums do not exist, i.e., when $(1 + \eta)\theta < 1$ and $\underline{h} = 0$.

population, P , total land supply, L , and real interest rates, r .

We assume that the distribution of abilities follows a (doubly-truncated) lognormal distribution.¹² We calibrate its parameters (μ, σ) to match two moments from the income distribution of São Paulo: the average income and the ratio of household income of the 80th percentile to that of the 20th percentile.¹³

The weight of housing consumption in the utility function, α , is set to 0.15. This value was drawn from Brazil's Consumer Expenditure Survey (IBGE (2004)). We set the weight of land inputs (γ) in the production of housing space to be equal to 0.25. The parameter γ can be rewritten as the share of land of the total output from the Cobb-Douglas production function for developers. Data drawn from construction companies listed on the São Paulo Stock Exchange show that the target value for land prices ranges from 25% to 35% of the total output of residential units. The property tax rate (η) equals 1%. In São Paulo, the property taxes range from 0.8% to 1.6% of the market value (housing price) of a housing unit. Although there is some variability, the median and mode rates for residential units were measured as 1%.¹⁴ Labour income tax (τ) is set to 22%. According to the Brazilian Revenue Service, the effective labour income tax reaches 12%, and there is an effective social security contribution of 10% on top of the labour income tax.

Parameters P and L are set to one. We also need to stipulate the interest rate r , for which we apply the national interest rate. Based on a 30-year average for Brazil, the interest rate is set to 4% per annum in real terms. Parameter v is the labour share in the production of the consumption good. We set it to 0.6, which is consistent with the estimates provided by Gollin (2002). Given r and v , we set productivity factor B such that the wage rate is equal to 1 in 2000.

Internal calibration. The three remaining parameters are protection costs, ψ , informality disutility, θ , and the minimum housing space, \underline{h} . We find the value of these parameters

¹²Parker (1999) argues that the lognormal distribution does not adequately fit the tails of the income distribution. We truncate 2.5% of the extreme observations from each tail to circumvent this issue.

¹³The data are from IBGE's PNAD. Online Appendix A presents the details.

¹⁴Note that η multiplies *housing rents* (R_H) in the model, so it is necessary to calculate the "effective" property tax for housing rents (instead of housing prices). For instance, 1% tax on housing prices is equivalent to a 25% tax on rents for a given interest rate of 4% per year.

by minimizing the distance between the model's predictions and the data. In particular, we use five moments in the minimum distance procedure: (i) the fraction of the population living in slums; (ii) the housing space Gini coefficient; (iii) the ratio of housing space of the 25th percentile to that of households of the 35th percentile; (iv) the ratio of housing space of the 45th percentile to that of the 55th percentile; and (v) the ratio of housing space of the 65th percentile to that of the 75th percentile. The goal is to obtain measures of centrality and dispersion from the housing space distribution. Let $\mathcal{L}(\xi)$ be the quadratic loss function to be minimised, let M_d be the vector of data moments, and let $M_m(\xi)$ be the corresponding vector of model moments. We want to find a $\hat{\xi}_{int} = (\psi, \theta, \underline{h})$ such that

$$\hat{\xi}_{int} = \arg \min_{\xi_{int}} \mathcal{L}(\xi) = \arg \min_{\xi_{int}} \frac{1}{N} [M_m(\xi) - M_d]' W_N \frac{1}{N} [M_m(\xi) - M_d]. \quad (23)$$

where $M_m(\xi) - M_d$ is the orthogonality condition and W_N is a positive semi-definite weighting matrix. We use the optimal weighting matrix W , which is given by the inverse of the variance-covariance matrix of data moments (S).¹⁵

Table 3 shows the parameter values. The only internally calibrated parameter whose magnitude has a direct economic interpretation is ψ . Parameter θ is a utility shifter, and \underline{h} is a catch up of all variables related to housing regulations in the formal sector. Field (2007) estimates that in Peru, having no legal claim to property is associated with a 14% reduction in total household working hours. Our estimate for São Paulo is that for each housing space unit, the cost of informality is approximately 24% of the disposable income. Thus, our estimate is in line with the empirical literature.

Table 4 shows the baseline economy for 2000. Our parameterisation is able to match the share of slums from the data. To compute housing space moments, we use information on the density of people per room (the census provides information on the number of rooms and on the number of people in each housing unit). A variable called “density per room” is used as a proxy for size. Four moments were created from this variable: a housing space

¹⁵The model is overidentified, as there are three parameters in ξ_{int} and five moments. As there are five moments, the orthogonality condition is a 5x1 matrix and the weighting matrix is a 5x5 matrix. In addition, we apply the following constraints to our estimation procedure: $\psi > 0$, $0 < \theta < 1$, $\theta(1+\eta) - 1 > 0$, which is equivalent to Assumption 1 and $\underline{h} > 0$.

Table 3: *Model Estimation: Parameters for the City of São Paulo*

External calibration			
Parameter	Description	Values	Source
α	Preferences: Weight of housing service	0.15	Expenditure Survey
γ	Housing: Weight of land input	0.25	Construction companies
η	Government: Property tax rate	0.25	São Paulo city council
τ	Government: Income tax rate	0.22	Ministry of finance
P	Urban population size	1	Normalised to one
L	Land supply	1	Normalised to one
r	Real interest rate	0.04	Central bank of Brazil
v	Labour share in production	0.6	Gollin (2002)
B	Total factor productivity	0.747	Normalised such that $w = 1$
μ	Average labour productivity	811.12	PNAD survey
σ	Income ratio: 80th percentile to 20th	17.85	PNAD survey
Internal calibration			
Parameter	Description	Point estimate	Source
ψ	Protection costs	0.24	Estimation
θ	Housing informality disutility	0.93	Estimation
\underline{h}	Minimum lot size regulation	0.40	Estimation

Table 4: *Moments: Data and Baseline Model*

Moments	Data	Model
Slums Share (%)	8.92	8.92
Gini housing space	0.39	0.48
Housing space: 25th/35th percentile	0.85	0.86
Housing space: 45th/55th percentile	0.80	0.78
Housing space: 65th/75th percentile	0.77	0.72

Notes. Data for São Paulo. Data moments calculated from Brazil's population censuses of 2000, which is the baseline year.

Gini index and three ratios of housing space for different percentiles. Table 4 shows that our parameterisation can match the empirical housing space distribution fairly well.

5 Quantitative Analysis

The quantitative exercises are performed in three parts. In the counterfactual analysis shown in Section 4.1, we quantify how much of the growth in slums is accounted for by changes in the distribution of income, (μ, σ) , and migration, P . For the policy experiments shown in Section 4.2, we investigate the quantitative effects of changes in the following policy parameters: land use regulation, \underline{h} , protection costs, ψ , informality disutility, θ , and property taxation, η . We close Section 4 with a welfare analysis.

5.1 Counterfactual Analysis

Our period of analysis is from 1980 to 2000.¹⁶ In Table 5, we decompose the contribution of the population size, average income and income inequality to slum growth. In counterfactual (1), we simulate the model assuming that the average income is the same as it is for 1980, while setting all other parameters to the 2000 value. For counterfactuals (2) and (3), we conduct the same experiment for income inequality and urban population, respectively. These exercises indicate that income inequality and rural-urban migration have the strongest impacts on slum formation from 1980 to 2000. Changes in *per capita* income explain a smaller fraction of changes in slum growth occurring in this period.

5.2 Policy Simulations

In this section, we analyse the effects of four policy experiments. First, we evaluate the impact of a reduction in real estate taxation η by 20%. Second, we simulate the effect of less strict man-made regulations through a 7% reduction in the value of \underline{h} . Third, we evaluate the effects of providing better basic infrastructure to informal settlements (by

¹⁶We do not use data for 2010 in the counterfactuals because slum data from the 2010 census are not directly comparable with those of other censuses (IBGE (2011)).

Table 5: *Decomposing the Variation in Slum Growth from 1980-2000*

	Fraction of population in slums in 2000
Baseline calibration	8.92%
Counterfactuals:	
1. Average income as in 1980, $\frac{\mu_{1980}}{\mu_{2000}} = 0.909$	10.91%
2. Income inequality as in 1980, $\frac{\sigma_{1980}}{\sigma_{2000}} = 0.756$	5.16%
3. Urban population as in 1980, $\frac{P_{1980}}{P_{2000}} = 0.704$	6.20%
4. Changing all	4.63%
Actual share in 1980	4.07%

Notes. To obtain the counterfactuals in 1980, we modify three parameters (urban population, average income, and income inequality) separately to their values in 1980. All other parameters are used as in 2000 (see Table 3).

setting $\theta = 1$). Finally, we simulate the impact of a land titling program (as represented by a decrease in protection costs ψ by 20%). In order to make the effects of each policy easier to compare, we change each parameter in such a way that the quantitative effects on slum formation are similar in absolute terms for each policy. We assume a balanced budget constraint from the local government such that g varies endogenously with each policy.

The effects of different policies on slums are shown in Table 6. The comparative statics indicate that a reduction in property taxation should decrease the number of slum dwellers (Proposition 2). According to Table 6, a reduction of property taxes η by 20% would decrease the informal housing presence by approximately 1.7 percentage points. The second simulation is associated with the impact of an overall reduction in regulations assuming that the entire city is now under less strict urban regulation. A 7% reduction in the “regulation bundle” decreases the share of slums by approximately 1.78 percentage point.

Table 6: *Policy Simulations: Macro and Average Welfare Effects*

	Fraction of population in slums in 2000	Average equivalent variation
Baseline calibration	8.92%	-
Policy simulations:		
1. Property tax, η : Reduction by 20%	7.23%	0.53%
2. MLS, \underline{h} : Reduction by 7%	7.14%	0.11%
3. Infrastructure upgrading, $\theta = 1$	10.71%	0.14%
4. Titling, ψ : Reduction by 20%	10.61%	0.11%

We then simulate the impact of two common policies on slums: infrastructure upgrading and titling. By infrastructure upgrading, we mean better infrastructure and public services.

Land titling means reducing the likelihood of eviction by giving property rights to slum dwellers. Policies on slums are typically classified under the umbrella of “slum upgrading” interventions. Usual interventions involve land titling, basic infrastructure provisions (e.g., water, electricity, and public lighting), the construction of facilities (schools, health posts), and home improvements. Even though slum upgrading interventions aim to provide both better infrastructure and ownership rights, we separate these two aspects to understand specific potential impacts.

The impact of providing infrastructure in slums is captured by increasing informal housing disutility (θ), but keeping the protection costs (ψ) at the same level as before. The provision of infrastructure (θ increases from 0.93 to 1) is associated with a 1.79 percentage increase in the number of slum dwellers. This means that upgrading interventions may have unintended impacts because it makes informal housing more “attractive”. We also decrease protection costs (ψ), while keeping disutility from informality, as it is in the baseline economy. The ex-post evaluation literature shows that giving property rights to the urban poor increases housing investment and housing prices (e.g., Lanjouw and Levy (2002) and Galiani and Scharfrodsky (2010)). Our simulations show that providing more security for informal housing should cause more people to live in informal settlements.

5.3 Welfare

We now evaluate the welfare impact of the policies described above. We measure welfare changes using the concept of equivalent variation (EV), which corresponds to the percentage change in income such that households should enjoy the same level of utility before and after a policy change.¹⁷

It should be stressed that the welfare analysis here does not necessarily account for the possible externalities created by infrastructure upgrading and land titling, such as a reduction of criminality and an improvement in sanitation and hygiene practices. The level of public goods provision, however, changes endogenously in the model with the policies

¹⁷In particular, EV is defined such that $V_b(\lambda(1+EV)) = V_a(\lambda)$, where V_b and V_a denote utility achieved under the baseline calibration and a counterfactual policy, respectively.

and may thus capture some of these effects. For instance, when policies decrease informal housing, revenue from property taxation rises and more public goods are available. In addition, we are not solving the problem of a benevolent social planner to determine the optimal government policy. Although interesting and relevant, this is not the goal of this paper.¹⁸

The average welfare effects of each policy are reported in the last column of Table 6. While a reduction of property taxation by 20% generates a similar impact on slum formation in absolute terms, it has the greatest effect on welfare. In fact, the consumption equivalent variation from reducing η by 20% is 0.53% relative to the baseline, nearly 4 to 5 times the gains attained by the other policy experiments. The intuition is that a change in property taxes affects all households living in formal housing and thus the measure of agents directly affected is large. In contrast, MLS relaxation, infrastructure upgrading and titling directly affect only a small subsample of households.

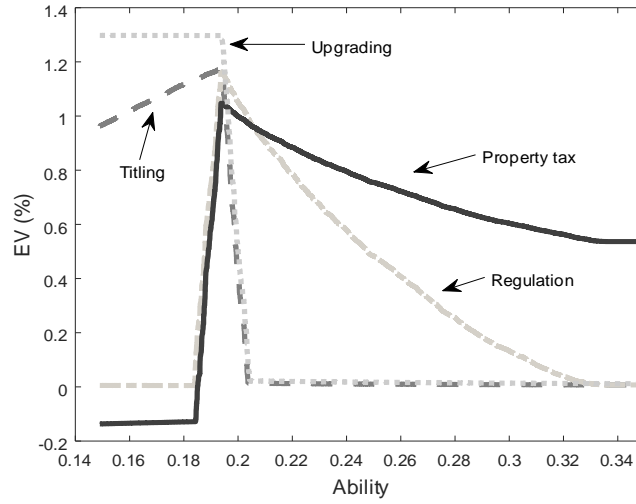
We also analyse the distributional welfare effects of each policy. In Figure 2, we show the consumption equivalent variation by ability. It can be seen that the reduction in property tax has a sizeable impact on welfare of high income agents, which increases by nearly 0.55%. The intuition is that, even though such households increase housing demand and, therefore, housing prices, the fall in property tax dominates and housing costs decrease. In contrast, low income households living in slums face higher housing costs after the policy change because they do not benefit from a lower property tax and, as a consequence, their welfare decreases by approximately 0.15%. Figure 2 also shows a spike in welfare for those individuals whose tenure choices are modified by the simulated policy.

The relaxation of MLS regulations presents a distinct welfare pattern. This policy directly affects individuals that are informal due to housing regulations and formal households whose the minimum lot size constraint is binding. It can be seen in Figure 2 that these households directly affected by a fall in h experience a welfare gain of nearly 1%. In contrast, low income households living in slums after the policy change and high ability

¹⁸We find that the provision of public goods g decreases only slightly after a reduction in property taxes, for instance. Property tax reduction induces formalization (which generates more revenue). In addition, labour taxes are an important component of total government revenues in our numerical exercises.

agents receive a tiny positive EV due to slightly lower housing prices.

Fig. 2: Individual-Level Equivalent Variation (EV)



Notes. The figures show the welfare impact measured by individual-level equivalent variations (EV) for each policy experiment.

As shown in Figure 2, infrastructure upgrading has a strong positive impact on households living in the informal housing sector (1.3% of income relative to the baseline). However, this policy has a very small impact on the rest of the population. The intuition is that an increase in θ causes some households to switch from formal to informal housing tenure, which reduces housing space consumption and housing prices and, as a consequence, generates a positive effect on the welfare of the other agents. The fall in housing demand happens because informal housing is still insecure and agents who switch do not have to comply with MLS regulations any longer. Nevertheless, given that the mass of individuals that change from formal to informal sector (as well as their housing demand) is small, this indirect effect on welfare is small.

Finally, Figure 2 also shows the effect of titling, which is represented by a decrease in ψ . This policy reduces the housing costs for informal housing and thus has a direct impact on low income agents living in slums. Welfare for this group increases by as much as 1.3%. In addition, titling also has an indirect effect on high income agents due to a fall in housing demand and prices. However, similarly to the infrastructure upgrading experiment, this general equilibrium effect acting through prices is very small.

In the companion online Appendix D, we relax two assumptions of our baseline model. First, we allow the housing rents of formal and informal housing to differ. In our model, the “effective cost” between formal and informal housing varies, but only one housing rent value is considered. Second, we endogenise the wage rate to account for general equilibrium effects acting through labour market adjustments; i.e., we add migration and a negatively sloped labour demand function. This is in line with a tradition in development economics of investigating the effects of rural-urban migration (Harris and Todaro (1970)). Relaxing these two assumptions implies that we cannot derive results analytically; therefore we focus only on the quantitative analysis. Our main findings are broadly unchanged when these two assumptions are relaxed.

6 Conclusion

In this paper, we investigate the main determinants of informal housing and evaluate the impacts of policies on slum formation. We construct a simple model of a city in which households can choose to live in formal or informal housing settlements. An important element of our theory corresponds to the protection costs. Our estimated parameter shows how a slum is associated with “time lost”: slum dwellers lose time to protect informal plots and recover from health problems related to an inadequate urban infrastructure, among other factors. These opportunity costs of living in slums are higher for richer households. The model is calibrated to be consistent with several statistics related to the urbanisation process of the city of São Paulo in Brazil. São Paulo is used as an illustration, but the model can be applied to other urban areas of developing countries.

Our results show that rural-urban migration and changes in the distribution of income explain much of the variation in slum growth from 1980 to 2000 in São Paulo. In particular, the results indicate that the rapid urbanisation process during this period had a considerable effect on slum growth. Our ex-ante policy evaluation analysis shows the following: (i) decreasing barriers to formalisation may largely slow down slum formation; (ii) incomplete slum upgrading interventions can have unintended impacts, as a sole improvement in slum

infrastructure can motivate more people to live in such areas; and (iii) the average welfare effect is mainly driven by strong impacts on the subset of the population directly affected by each policy.

Tiago Cavalcanti, University of Cambridge and Sao Paulo School of Economics - FGV

Daniel Da Mata, Institute for Applied Economic Research - IPEA and Sao Paulo School of Economics - FGV

Marcelo Santos, Insper Institute of Education and Research

References

- Amaral, P.S. and Quintin, E. (2006). ‘A competitive model of the informal sector’, *Journal of Monetary Economics*, vol. 53(7), pp. 1541 – 1553.
- Antunes, A.R. and Cavalcanti, T. (2007). ‘Start up costs, limited enforcement, and the hidden economy’, *European Economic Review*, vol. 51(1), pp. 203 – 224.
- Baum-Snow, N. (2007). ‘Suburbanization and transportation in the monocentric model’, *Journal of Urban Economics*, vol. 62(3), pp. 405–423.
- Brinkman, J., Coen-Pirani, D. and Sieg, H. (2015). ‘Firm dynamics in an urban economy’, *International Economic Review*, vol. 56(4), pp. 1135–1164.
- Brueckner, J.K. (2013). ‘Urban squatting with rent-seeking organizers’, *Regional Science and Urban Economics*, vol. 43(4), pp. 561 – 569.
- Brueckner, J.K. and Selod, H. (2009). ‘A theory of urban squatting and land-tenure formalization in developing countries’, *American Economic Journal: Economic Policy*, vol. 1(1), pp. 28–51.
- Cai, Y., Selod, H. and Steinbuks, J. (2018). ‘Urbanization and land property rights’, *Regional Science and Urban Economics*, vol. 70, pp. 246 – 257.
- Chatterjee, S. and Carlino, G.A. (2001). ‘Aggregate metropolitan employment growth and the deconcentration of metropolitan employment’, *Journal of Monetary Economics*, vol. 48(3), pp. 549–583.
- Da Mata, D., Deichmann, U., Henderson, J., Lall, S. and Wang, H. (2007). ‘Determinants of city growth in Brazil’, *Journal of Urban Economics*, vol. 62(2), pp. 252–272.
- Da Mata, D., Lall, S.V. and Wang, H. (2008). ‘Favelas e dinâmica das cidades brasileiras’, in (A. Carvalho, C. Oliveira, J. Mota and M. Piancastelli, eds.), *Ensaio de Economia Regional e Urbana*, pp. 47–64, Brasília: IPEA.
- Di Tella, R., Galiani, S. and Schargrodsy, E. (2007). ‘The formation of beliefs: Evidence from the allocation of land titles to squatters’, *The Quarterly Journal of Economics*, vol. 122(1), pp. 209–241.

Field, E. (2005). 'Property rights and investment in urban slums', *Journal of the European Economic Association*, vol. 3(2-3), pp. 279–290.

Field, E. (2007). 'Entitled to work: Urban property rights and labor supply in peru', *Quarterly Journal of Economics*, vol. 122(4), pp. 1561 – 1602.

Friedman, J., Jimenez, E. and Mayo, S.K. (1988). 'The demand for tenure security in developing countries', *Journal of Development Economics*, vol. 29(2), pp. 185 – 198.

Galiani, S., Gertler, P.J., Undurraga, R., Cooper, R., Martinez, S. and Ross, A. (2017). 'Shelter from the storm: Upgrading housing infrastructure in latin american slums', *Journal of Urban Economics*, vol. 98(C), pp. 187 – 213.

Galiani, S. and Schargrodsky, E. (2010). 'Property rights for the poor: Effects of land titling', *Journal of Public Economics*, vol. 94(910), pp. 700 – 729.

Gollin, D. (2002). 'Getting income shares right', *Journal of Political Economy*, vol. 110(2), pp. 458–474.

Harris, J.R. and Todaro, M.P. (1970). 'Migration, unemployment and development: A two-sector analysis', *American Economic Review*, vol. 60(1), pp. 126–142.

IBGE (2004). *Pesquisa de Orçamentos Familiares: 1986-1987, 1995-1996, and 2002-2003*, Rio de Janeiro, RJ: Instituto Brasileiro de Geografia e Estatística - IBGE.

IBGE (2011). *Censo 2010: Aglomerados Subnormais, Primeiros Resultados*, Rio de Janeiro, RJ: Instituto Brasileiro de Geografia e Estatística - IBGE.

Jimenez, E. (1985). 'Urban squatting and community organization in developing countries', *Journal of Public Economics*, vol. 27(1), pp. 69–92.

Kapoor, M. and le Blanc, D. (2008). 'Measuring risk on investment in informal (illegal) housing: Theory and evidence from pune, india', *Regional Science and Urban Economics*, vol. 38(4), pp. 311–329.

Kopeccky, K. and Suen, R. (2010). 'A quantitative analysis of suburbanization and the diffusion of the automobile', *International Economic Review*, vol. 51(4), pp. 1003–1037.

Lanjouw, J.O. and Levy, P.I. (2002). 'Untitled: A study of formal and informal property

rights in urban ecuador', *The Economic Journal*, vol. 112(482), pp. 986–1019.

Marx, B., Stoker, T. and Suri, T. (2013). 'The economics of slums in the developing world', *Journal of Economic Perspectives*, vol. 27(4), pp. 187–210, doi:10.1257/jep.27.4.187.

Michaels, G., Rauch, F. and Redding, S.J. (2012). 'Urbanization and structural transformation', *The Quarterly Journal of Economics*, vol. 127(2), pp. 535–586.

Parker, S.C. (1999). 'The generalised beta as a model for the distribution of earnings', *Economics Letters*, vol. 62(2), pp. 197 – 200.

UN-Habitat (2016). *Urbanization and Development: Emerging Futures / United Nations Human Settlements Programme*, Earthscan Publications, London.