Published in final edited form as:
J Urban Econ. 2011 May 1; 69(3): 253–272. doi:10.1016/j.jue.2010.09.006.

Exclusionary policies in urban development: Under-servicing migrant households in Brazilian cities

Leo Feler[a] and **J. Vernon Henderson**[b,*]

[a]Johns Hopkins University, School of Advanced International Studies, Department of International Economics, 1717 Massachusetts Avenue NW, 7th Floor, Washington, DC 20036, United States

[b]Brown University, Department of Economics, Box B, Providence, RI 02912, United States

Abstract

Localities in developed countries often enact regulations to deter low-income households from moving in. In developing countries, such restrictions lead to the emergence of informal housing sectors. To deter low-income migrants, localities in developing countries withhold public services to the informal housing sector. Using a large sample of Brazilian localities, we examine migration and exclusion, focusing on the public provision of water to small houses where low-income migrants are likely to live. Withholding water connections reduces the locality growth rate, particularly of low-education households. In terms of service provision, during dictatorship in Brazil, we find evidence of strategic exclusion, where localities appear to withhold services to deter in-migration. We also find evidence of strategic interactions among localities within metro areas in their setting of service levels: if one locality provides more services to migrant households, other localities respond by withholding service.

Keywords

Urban slums; Urban growth; Urban services; Exclusion

> Macae, Brazil (Reuters, March 19, 2008) – Thin black water hoses snake across the ground all over Nova Holanda, a workers' neighborhood on the edge of the Brazilian boom town of Macae. Oil and gas may have brought riches to the former fishing village but water is almost as precious. Over the years, thousands of job seekers have flocked to the area, building homes wherever there's a patch of ground. But Nova Holanda has only one standpipe linked to the city's main system – forcing people to install a makeshift system of their own with hoses…. [Santos] and many other people installed the hoses so they can sell water from the city system at a price of about 10 reais ($6) for 100 liters. The rough conditions make Nova Holanda a hostile, dirty and violent place…. Macae represents the country well, with its wealth concentrated in a very small elite. "We cannot end those infrastructure flaws that have lasted decades…. Eventually it will be finished, may be in thirty years." (Romulo Campos, city spokesman).

1. Introduction

In developing countries, during periods of rapid urbanization, urban areas often house significant portions of their population in slums or informal housing. For Brazil, this is

*Corresponding author. Fax: +1 401 863 1970, lfeler@jhu.edu (L. Feler), vernon_henderson@brown.edu (J.V. Henderson).

JEL classification: D7, H7, J6, O15, O54, R5

illustrated by the emergence of its now infamous favelas. Similar settlements are also a growing phenomenon in South and Southeast Asia, Sub-Saharan Africa, and in large Chinese urban areas. Informal housing sectors are usually characterized by low housing quality, varying degrees of tenure insecurity, and perhaps more critically, they tend to be cut-off from basic urban services, which makes living conditions expensive, unpleasant, and unhealthy.

This paper seeks to examine why such areas exist. Informal housing sectors that lack basic services might arise because cities do not have sufficient resources or sufficiently developed institutions to provide basic services. In many cases, however, we observe the proliferation of informal housing in some neighborhoods of the largest and wealthiest cities of developing countries, which ostensibly have both the institutions and means to provide basic services to the informal sector. Sometimes it is argued that the low-income residents in the informal sector cannot afford such services. For the service we focus on, water, we make two points. First, the private alternative to public servicing is much more costly, so residents can better afford the public service than the private one they must buy. Second, hedonic regressions suggest residents value these services. Rather than the issue being one of affordability or lack of institutions, we hypothesize there is a strategic element to withholding services from informal sector households. By withholding services, cities can perpetuate bad living conditions for informal sector residents and discourage further in-migration of the types of people likely to occupy these areas.

In developed countries, attempts at exclusion take a different form. Discouraging entry into the most desirable cities involves formal sector housing restrictions—such as minimum lot size zoning, building height restrictions, and frontage requirements— which drive up the total cost of a house and make these locations prohibitively expensive for poorer in-migrants. “Superstar” cities which restrict development experience much slower growth in their number of households and increases in their share of the population from high-income groups, as high-income migrants outbid and drive out poorer residents (Gyourko et al., 2006). But such restrictions can only be effective when institutions prohibit the emergence of an informal housing sector. If establishing formal sector housing restrictions results in the growth of an informal sector, as is the case in many developing countries, then one approach to limit entry into higher income cities might be to make living conditions in the informal sector more difficult.

In China today, this strategic component of withholding services is explicitly articulated in policy design (Cai, 2006). Forcing the vast majority of migrants into poorly serviced informal settlements is intended to restrain rural in-migration to China's largest urban areas. While explicit articulation of exclusionary motivations for withholding public services is less feasible politically in countries like Brazil, we present data indicating such policies were nevertheless adopted.

We focus on the provision of public water in the period 1980– 2000 across Brazilian localities (spatial units roughly equivalent to US counties) to the types of houses in which poor migrants are likely to live. We focus on Brazil because it is a large developing country with a decentralized system of government, and we focus on public water provision for several reasons. First, water is an essential good, and in the absence of public provision, residents must procure it through other, usually more expensive and inconvenient means, such as purchasing it from water delivery trucks, buying bagged or bottled water, or buying it from other residents who pump it from often polluted rivers and streams. Second, unlike electricity, it is very difficult for residents to illegally connect and steal from the public water distribution system (which would involve digging-up streets and laying pipes). Third, even when water is managed by a state or regional entity, decisions on which neighborhoods to service are made by the municipality, so servicing is very much a local decision. Fourth, during the period we analyze, all localities in our sample have a public water distribution system in place, which is not the case for other services such as sewerage. Thus for public water we have a decision on whether or not to service households, in a context where a system is in place for households to potentially be serviced. Finally, we can accurately measure, using census data, which houses have public water connections and which do not, and we can define houses as being in the informal sector based on their servicing. Other measures of informality, as we will note later, such as whether neighborhoods are “irregular” or whether owners possess land title, are both less precise and less relevant in the context of Brazil.

In this paper, we examine, first, the impact of withholding public water services to the houses most likely to be occupied by poor migrants on a locality's subsequent growth in number of households and social composition, and second, what determines a locality's provision of services to housing likely to be occupied by poor migrants. An additional motivation for this study is that the proliferation of unserviced housing has effects beyond possibly restraining in-migration, such as generating inequality in living conditions and unhealthy neighborhoods with large negative externalities. The resulting negative externalities may also affect the location decisions of those who live in nearby formal sector housing.

Our findings suggest that as the share of small houses (likely to be occupied by poor migrants) connected to public water increases, so does a locality's subsequent growth rate of urban households. When we decompose this effect by the education level of the household head, we find that servicing a larger share of small houses leads to faster growth of low-education households. It also has a positive but noisy effect on the growth of high-education households, which is potentially due to reduced negative externalities from having fewer under-serviced neighborhoods affecting quality-of-life nearby. We find that larger localities service a larger share of small houses (a scale effect), as do richer localities (a wealth effect); and we explore non-linearities in scale and income, looking for diminishing effects. A key result is that the interaction of being rich and large is strongly negative. We interpret this interaction as evidence of “exclusion,” whereby “superstar” localities service relatively fewer small houses. This helps explain why the largest and wealthiest localities experience both slower population growth and relative increases in the share of their population from high-income groups.

In examining the determinants of service provision, we also test whether localities within a metro area behave strategically with one another in setting service levels. We find evidence of negative strategic interactions: localities appear to respond to increases in servicing in other localities by withholding their own service. This is the opposite of a race to the bottom or top, where localities uniformly increase or withhold servicing. As some localities within a metro area increase servicing, this attracts migrants to the metro area overall. But this increased migration induces other localities to withhold servicing in order to deflect incoming migrants away from their own localities.

Overall, we find evidence of strategic under-provision of public water service to the types of houses most likely to be occupied by poor migrants in Brazil's localities. This under-provision appears to function as a deterrent to further entry of poor migrants, and it can explain why Brazil's “superstar” localities, which have the wealth and scale to expand servicing, also have informal neighborhoods that lack servicing. If these localities were to provide servicing to informal neighborhoods, and implicitly grant these areas formality, they would encourage further entry of poor migrants into their jurisdictions and experience further congestion and dissipation of the amenities that make them desirable places to live.

The paper proceeds as follows. Section 2 discusses data, trends, and relevant literature concerning exclusion in Brazil and in developed countries. Section 3 develops a conceptual framework to inform econometric specifications. Section 4 estimates the impact of providing public water services to small houses on a locality's subsequent growth in number of households and social composition. Section 5 analyzes how a locality sets its public water servicing and Section 6 concludes.

2. Urbanization and public service provision in Brazil

This paper focuses on the post-1980 time period in Brazil. We first examine the population response to public water service across localities, specifically how public water provision in 1991 to the houses in which low-skilled migrants are likely to live affects a locality's growth in number of urban households and social composition between 1991 and 2000. After establishing that public water provision affects in-migration, we turn to the determinants of public water provision. We focus on provision during the 1980s under dictatorship.

This timing turns out to be convenient. First, in terms of an identification strategy, the 1980s mark the end of Brazil's urbanization and state involvement in the economy, which placed industry in the biggest cities in the post World War II period (Baer, 2001, pp. 387–390). Industrial development, which had focused on Sao Paulo and Rio de Janeiro starts to decentralize in the 1970s, with substantial and on-going industrialization of hinterland localities. This decentralization is facilitated by inter-locality investments in transportation, telecommunications, and finance, as well as agricultural developments in the Northeast of Brazil (Da Mata et al., 2005). By the 1990s, these adjustments are largely complete. This change in urbanization and industrialization patterns reflects changes in the underlying drivers of locality growth from the pre-1970 period to the 1990s. We use this change in the determinants of locality growth as part of the identification strategy in Section 4.

Second, the transition from dictatorship to democracy during the late 1980s, allows us to assess the determinants of public water provision under different regimes. Our working assumption is that during the 1980s, exclusionary under-servicing of informal neighborhoods by localities was possible and politically feasible. Elitist dominated localities could legitimately withhold services from the informal sector. However, with full democratization at the national level in 1988, the introduction of mandatory voting laws, and reforms promoting the regularization of informal sector housing, such exclusionary behavior on the part of localities was no so longer politically feasible. We observe this in the data: while we find evidence of exclusionary under-servicing in the 1980s, by 2000, this is no longer the case. The actions of localities in Brazil in the 1980s and their impacts through the 1990s provide lessons for countries today at a lower level of urbanization that are following a similar development path as Brazil.

2.1. Data

We use population census data for Brazil from 1970, 1980, 1991, and 2000. These contain a variety of information on housing size, tenure mode, and servicing of houses, as well as basic socioeconomic information covering education, income, family structure, and migration. We also have information on geographic and fiscal indicators.

Local governments in Brazil are municipalities (municipios), units equivalent to counties in the United States. Because some municipalities split over time and some are annexed by other municipalities, following Da Mata et al. (2005), we combine municipalities into common denominator ones, which we refer to as “localities”, or more formally, Minimum Comparable Areas (MCAs). These localities are constant spatial units during the time period we analyze. We focus on localities that are at least 50% urbanized by 1991 and that are located in larger metro areas (approximately equivalent to Metropolitan Statistical Areas in the United States) with multiple localities. What fraction of a locality is "urban" depends on density, nature of employment in different neighborhoods, and commuting patterns, and the definition, provided by the census, remains fairly constant over time. Metropolitan areas are based on the definitions in IPEA, IBGE, and UNICAMP (2002).

For our base sample, we exclude all localities that form their own isolated metro area since we are interested in how localities within a metro area might interact in the provision of services and since we wish to control for metro area fixed effects. This leaves us with a base sample of 327 localities in 54 metro areas. We will perform robustness checks with different samples, where we exclude localities that are a recombination of municipalities and where we exclude localities in the Rio de Janeiro and Sao Paulo metro areas.

2.2. Urban growth and stratification

In the spatial development of Brazil, metro areas experience mostly parallel growth from 1980 to 2000, as a number of theories predict (Black and Henderson, 1999; Gabaix, 1999). While metro areas grow mostly in parallel, with knowledge accumulation and improved education levels being the main drivers of metro area growth (Da Mata et al., 2007), this is not the case for localities within metro areas. Localities within metro areas experience strong mean reversion: more populous (larger) localities grow more slowly than less populous ones, as shown in Fig. 1. Also in Fig. 1, points marked with a dark square denote the top 20% of localities sorted on the basis of median income. These richer localities are generally the larger and slower growing ones.

With population growth in a metro area, old localities fill up and become crowded, and new localities develop. With economic development, metro areas spread out, fueled by declining commuting costs and transportation improvements that make central locality locations less valuable. This movement of migrants into different localities within metro areas is the variation we utilize in our empirical work. In Brazil, as in much of the world, the rich live more in central localities and the poor live more in suburbs. Much of the exclusion we observe is therefore of over-crowded "superstar" central localities and richer and larger suburbs deflecting poor migrants into low-income, suburban localities within the metro area.

Finally, large and wealthy localities tend to become even wealthier over time. In Fig. 2, we plot the share of a locality's households that are among Brazil's 10% richest households in 1980 and in 2000. Those localities that have a disproportionate share of the rich in 1980 are almost universally above the 45-degree line, with an even greater share of Brazil's richest households by 2000. These localities also tend to be among the 20% most populous of the localities in our sample.

The data seem to suggest an element of the superstar cities story occurring in Brazil: large, wealthy localities are growing more slowly and becoming disproportionately richer over time. But unlike the Gyourko et al. (2006) version of this story for the United States, in Brazil where an informal housing sector can exist alongside a formal housing sector, formal sector housing restrictions alone are not sufficient to constrain housing supply and produce the slower population growth and higher social stratification we observe in Brazil's larger and richer localities.

2.3. The informal sector and public infrastructure

There are several ways to define and identify the informal housing sector as discussed in the literature (Biderman, 2007; Dowall, 2006). In the census, there is a question filled out by census workers on whether people live in irregular settlements. Irregularity in this context captures whether streets are straight or crooked and whether houses are properly numbered in a neighborhood configuration of housing, not whether houses are serviced or owners have formal title. Thus, irregularity differs from informality, and less than 5% of households are considered irregular.

Economists typically prefer to define informality based on ownership rights. In the 1991 and 2000 census, there is a question for home owners on whether they have title to their land. In 1991, about 9% of urban households living in owner-occupied housing in our localities report that they do not have land title. The number seems small compared to estimates in the literature; the belief is that many households without true title answer yes to having title because they do not feel insecure about their holdings. Home ownership is easily transferable even without formal land ownership, and evictions from areas where residents lack land title do not happen. Most non-titled houses are not in *favelas* which are "squatter" type settlements, but are in *loteamontos*. Even in *favelas*, evictions are rare. *Loteamentos* are developments where houses are built on land legally acquired by the developer but where the houses typically violate a 1979 national law imposing a minimum lot size of 125 square meters. Because the housing is in violation of this law, formal title cannot be obtained; but transferability of ownership is not impeded. The issue is that violation of zoning laws and lack of titling in a neighborhood means such settlements are "illegal" and localities in the 1980s were not required to service them.

Thus a different approach to defining informality is based on lack of public infrastructure provision and is used by researchers on Brazil like Dowall (2006). Dowall suggests a key element is a central water connection, where in 1991, about 17% of urban households were not connected, on average, across our sample of localities. A stronger criterion is to impose "full service": electricity (virtually universal in 1991), a central water connection, and a central sewer connection. In 1991, about 65% of households do not have full service, on average, across our sample of localities, which is primarily because they lack sewerage. In part, this is because many of the localities in our sample do not have public sewer systems in place: only 187 of the 327 localities in our sample had more than 10% of houses with "full service" as of 1991. In contrast, virtually all our localities had wide-spread public water distribution systems by 1991, with at least 10% of houses serviced. In neighborhoods that lack central sewer connections, residents must rely on septic systems—although these are not functional in high-density areas typically associated with informal settlements—or they empty their sewage directly into open ditches or streams. For water, one potential private alternative is to dig wells, although this is usually feasible only for wealthier households given the expense of digging and maintaining wells, and in low-density areas where there is less depletion and more exposed surface area so that water tables can be replenished by rainfall. In the high-density areas typically associated with informal urban settlements, the private alternatives at the margin are to use a public stand-pipe and haul water for some distance, extend hoses into nearby but often polluted rivers and streams and pump water into homes, subscribe to water delivery truck services, or purchase bottled or bagged water. These are difficult, dirty, and expensive alternatives to central water provision.

As a preliminary check that central water connections are highly valued, we examine willingness-to-pay for water connections using simple hedonic regressions for renters in the central localities of Sao Paulo and Rio de Janeiro. Hedonic regressions reveal willingness-to-pay within a locality for infrastructure connections, for those on the margin between choosing a serviced versus unserviced rental unit. The results are presented in Appendix A. Controlling for neighborhood fixed effects and an extensive list of house characteristics, we find the marginal renter is willing to pay 9% more for a rental unit with a central water connection in both Sao Paulo and Rio de Janeiro, net of any premium for indoor plumbing. When we restrict the sample to the poorest 20% of households in these localities, we find the marginal willingness-to-pay for a central water connection is 9% in Sao Paulo and 12% in

Rio de Janeiro. We take this as suggestive evidence that having a central water connection is highly valued, even when the marginal renter is low-income.

In summary, as our base measure of exclusion, we focus on the provision of public water connections to the types of homes likely to be occupied by low-income migrants. Other measures of exclusion, such as whether homes are in irregular neighborhoods, whether owners have land title, or whether homes are connected to sewerage, have less clear interpretations. Also, it seems that denying a proper neighborhood configuration or land title does not serve as an effective exclusionary mechanism. In the literature, lack of a central water connection and clean water (and the resulting illnesses from water-borne diseases) are key issues for residents of informal sector housing in Brazil (see Scheper-Hughes, 1993, pp. 65–97; World Bank, 2001). Furthermore, virtually all localities in our sample have a public water distribution system in place, and the decision to service neighborhoods is effectively determined by the locality (i.e., the *municipio*).[1] Withholding water connections to poor neighborhoods is a way of influencing the location decisions of low-income households. In some other developing countries (China and India, for example), exclusionary policies can target individuals based on personal characteristics, such as migrant status or, historically, their caste. In Brazil, exclusionary policies cannot single-out individuals, but can instead target the houses and neighborhoods where certain types of individuals are most likely to live.

2.4. Provision of infrastructure services

We explore dimensions of public water servicing in Table 1. In Panel A, we have the share of urban houses with a connection to the locality's public water system for localities that are at least 50% urbanized by decade. The sample expands in each decade as localities become more urbanized and surpass the 50% threshold. There is also rapid expansion in the share of urban households with a connection to the locality's public water system, from 50% in 1970, to 69%, 83%, and 89% in 1980, 1991, and 2000. This is the average share of urban households connected to the public water system in each locality. When we weight this average by the number of urban households in a locality, obtaining the average share of urban households serviced for all localities, we see that shares are systematically higher: more populous localities provide public water connections to a greater share of households, but the differences in the unweighted and weighted averages become smaller over time, suggesting that less populous localities catch-up to more populous ones in their servicing, particularly after 1991.

Looking more closely at the average share of urban houses serviced in 1991, 74% of those who own the house but not the land (who do not have land title) have connections to the public water system versus 85% for those who do have title. Next we look at servicing of small versus large houses. Small houses are defined on the basis of total number of rooms, with the cut-off being approximately the smallest 15% of houses by size. Similarly, large houses are those for which the total number of rooms place them in the largest 15% of

[1]Prior to 1968, the responsibility for water supply and sanitation was entirely municipal, with service providers being municipal water and drainage companies, each of them with different financial and administrative structures. Between 1968 and 1986 (during dictatorship), the National Water Supply and Sanitation Plan (known as “PLANASA”) created 27 state-owned water and sanitation companies, with dedicated lines of funding from the National Housing Bank (BNH) and the Employment Guarantee Fund (FGTS), and with the responsibility for construction, operation, and maintenance of water and sanitation infrastructure. But since the Brazilian Constitution had already established that the power to grant licenses for public water and sanitation services belonged to the municipalities, the state-owned water and sanitation companies needed to obtain concessions from the municipalities in order to operate locally. About 20% of municipalities during this period chose to continue providing services entirely on their own through municipal companies. After democratization and decline in federal funding, many states devolved services back to the local level. For the purposes of this paper, what is important is that municipalities have the ability throughout the period we analyze to determine which neighborhoods are prioritized for servicing, either through the management of their own municipal water and sanitation companies or through the granting of concessions and licenses to state-level companies (see Arretche, 2001; Heller, 2009).

houses, or above the 85th percentile.[2] Seventy-four percent of small houses have public water connections versus 89% for large houses. And lastly, 76% of the poorest quintile of households have public water connections versus 89% for the richest quintile.

Panel B of Table 1 shows similar results for a constant sample of 185 localities that were at least 50% urbanized by 1970. Analyzing now only on one margin—of increasing service but not adding new localities to the sample—we see that localities dramatically increase the share of urban households they service, from a mean of 50% in 1970 to 92% by 2000. Even when we look at localities that were less than 50% urbanized in 1970 (in Panel C), we see that these newer localities also have rapid expansion but are at lower service levels. These numbers suggest that localities can quite rapidly expand their central water systems, and the fact that some houses remain unconnected to public water systems by 1991 and 2000 may thus reflect attempts to deter further in-migration.

Localities presumably withhold public water connections to discourage in-migration of certain types of individuals. To implement this, since in Brazil localities cannot discriminate on the basis of migration (or race or income) status, they must discriminate by the types of houses migrants are likely to occupy. Migrants, and especially low-skilled, rural, and low-income migrants, disproportionately occupy small houses and are underrepresented in larger houses. In 1991, about 18% of households lived in small houses and 17% lived in large houses. But 29% of migrants (those who moved to the locality within the past five years), 45% of migrants from rural areas, 52% of poor migrants (those in the bottom quintile of the locality's household income distribution), and 41% of low-education migrants (those with less than a primary school education) lived in small houses. In contrast, 16% of non-migrants lived in small houses in 1991. Thus, targeting the smallest houses appears to be a mechanism to discriminate against migrants, and especially those who are from rural areas, poor, or have little education.

In Table 2, we examine the share of small versus large houses that are connected to the public water system for the full sample of 327 localities. Only 28% of small houses had public water connections in 1970 compared to 61% of large houses; in 1980, these shares are 49% and 78% for small and large houses, respectively; by 1991, they are 74% and 89%, and by 2000, they are 84% and 91%. The differences in servicing diminish considerably over time, especially after democratization in the mid-1980s. In Sections 3 and 4, we use the provision of public water connections to small houses as our basic exclusionary measure, representing the quality of infrastructure that incoming migrants, especially rural, poor, or low-education ones, might expect when choosing to live in a particular locality.

3. Conceptualizing exclusionary behavior

This section develops a model examining service levels and locality population growth to motivate key aspects of the empirical formulations. The empirical work focuses on variations across localities within metro areas for identification. For much of the empirics,

---

[2]We define small and large houses so that they approximately account for the bottom 15% and top 15% of all houses by total number of rooms. In 1970 and 1980, small houses are those with less than two rooms total (the bottom 14% and 13%, respectively, of all houses), and in 1991 and 2000, they are houses with less than three rooms (the bottom 18% and 16%, respectively). Large houses are those with more than seven total rooms in 1970 (the top 17% of all houses), and with more than eight total rooms in 1980, 1991, and 2000 (the top 14%, 17%, and 18%, respectively). We focus on total rooms instead of total bedrooms for a few reasons: first, there is greater variation in total rooms as opposed to total bedrooms, as might be expected, and second, "bedrooms" are defined as any room where residents sleep, and so is more ambiguous than a definition based on total rooms. While a possible concern is that houses with public water connections might have bathrooms and therefore have more total rooms, the literature (Scheper-Hughes, 1993) suggests that the main benefit of a public water connection for poor households is for cooking and drinking. Many houses have indoor plumbing but no bathrooms. In 1991, when the relevant census data allow distinctions, we observe that 40% of one-room houses have indoor plumbing (and hence no separate bathroom) and more generally that 21% of small houses that have indoor plumbing do not have a bathroom.

this means we can ignore the determination of metro area characteristics, which are captured by metro area fixed effects. For example, we assume workers in all localities in a metro area participate in the same overall metro area labor market. Then, conditional on total metro area size, people's choice of a locality within a metro area does not affect their wage incomes.

We formulate the basic problem much like the welfare competition literature in the United States (Wildasin, 1991), where within a region, localities choose policies in the face of a potential influx of migrants. In our case, the policy is the servicing of small houses typically occupied by migrants. The metro area faces a supply of in-migrants, which will be split across the localities of the metro area depending on the living conditions in these localities. Incumbent residents may want fewer migrants to their own locality because of congestion and other population externalities within the locality, as well as potential prejudices against having lower-income neighbors. This serves as a force to withhold servicing to migrant households. However, forces for incumbent residents to improve servicing of migrant households also exist. These include altruism and externalities, where inadequate servicing of migrant neighborhoods may generate negative externalities for incumbents and for migrants alike. Also, for economic growth reasons, incumbent residents of a locality may want more migrants to the overall metro area, but not necessarily to their own locality.

To develop these trade-offs, we start by specifying the preferences and demand functions of migrants depending on whether they are serviced or not. Then we look at equilibrium in the locality housing market and equilibrium in the flow of migrants to the metro area, as well as the distribution of migrants across localities. Based on this information, on the service levels in other localities, and on the characteristics of the own locality's incumbent-resident population, each locality chooses a level of servicing. While we model the equilibrium in a one-shot context of a locality facing a potential influx of migrants, throughout we emphasize that localities have a past history that influences their current characteristics.

All migrants are assumed to live in the informal housing sector while incumbent residents who make policy decisions live in the formal sector. Based on policy decisions of the locality, some migrants live in informal sector neighborhoods where the locality publicly provides piped water at a unit cost c_0 (e.g., the cost of metered water). Other migrants live in unserviced informal sector neighborhoods, where they must privately procure services at a higher unit cost, $c > c_0$ (e.g., the cost of water purchased from water delivery trucks). By living in the informal sector, we assume all migrants escape locally set property taxes, paid by incumbents.

3.1. Equilibrium within the informal sector

Migrants entering the informal sector of a locality who choose to live in serviced neighborhoods have quasi-indirect utility functions and housing demand functions of the form

$$U_0 = U_0(w(\overline{N}+\overline{L}), p_0, c_0, L_0, L, Z), \tag{1a}$$

$$h_0 = h_0(w(\overline{N}+\overline{L}), p_0, c_0, L_0, L, Z). \tag{1b}$$

Disposable income of migrants is $w(\bar{N}+\bar{L})$, which could vary with overall metro area scale as measured by total incumbents in the urban area, $\bar{N}$, plus total migrants to the urban area, $\bar{L}$. These represent scale externalities in overall labor markets of the metro area. In the locality, p_0 is the price of housing in serviced neighborhoods and c_0 is the unit cost of water.

Of the total migrants, L, to the locality, a number, L_0, are serviced, potentially a positive externality (relative to no servicing) for migrants as well as for incumbent residents, as will be discussed further below. That is, the utility of migrants is increasing in L_0 holding all else constant (i.e., $U_{L0} \geq 0$ with $L_0 \leq L$). For example, more migrant households serviced means fewer people using and disposing of untreated water, which could result in health externalities. The number of migrant households serviced is a policy variable influenced by the locality. The endogenous number of migrants from the metro area pool, $\bar{L}$, to the locality itself is L, which results in congestion or other negative externalities for migrants and incumbents. For migrants, this means that $U_L \leq 0$. Z is a set of locality characteristics, such as the size and income level of the incumbent population.

Migrants to the informal sector of the locality who live in unserviced neighborhoods have indirect utility and housing demand functions of the form

$$U=U(w(\overline{N}+\overline{L}), p, c, L_0, L, Z), \quad (2a)$$

$$h=h(w(\overline{N}+\overline{L}), p, c, L_0, L, Z). \quad (2b)$$

For the same housing price in a serviced (p_0) versus unserviced (p) neighborhood, $U_0 > U$, because $c > c_0$. An equilibrium allocation requires equalized utility across the two types of neighborhoods within a locality, which implies $p_0 > p$. Equating indirect utility in Eqs. (1a) and (2a) for the two types of neighborhoods, we know

$$p_0=p_0(p, L_0, c, c_0, L, Z;\overline{N}+\overline{L}). \quad (3)$$

Eq. (3) underlies the hedonic regressions reported in Appendix A, examining within locality differences in relative rents based on type of service received by a house. Across localities, overall differences in servicing and other conditions will be reflected in absolute price differences in both p_0 and p.

Migration to a locality is governed by two conditions: demand must equal supply in the locality housing market, and utility must be equal for all migrants across localities within the metro area. We turn to this next.

3.1.1. Housing demand equals housing supply, and the supply of migrants—

For housing demand equaling housing supply, we assume housing supply for migrants to the informal sector of a locality is given by $H^s(A, p)$, where A describes supply conditions in the locality, based upon vacant land availability and the cost of bringing extra land into production of housing services. The supply margin in a locality is the unserviced sector with price p. Hence the supply specification $H^s(A, p)$. Summing the individual housing demands of the L_0 serviced people in Eq. (1b) and of the $L - L_0$ unserviced people in (2), and using Eq. (3) for p_0, we have

$$L_0h(w(\cdot), p_0(\cdot), c_0, L_0, L, Z)+(L-L_0)h(w(\cdot), p, c, L_0, L, Z)=H^s(A, p). \quad (4)$$

The final piece for internal locality equilibrium concerns the supply of migrants to the locality. Localities within a metro area share migrants, whose total supply is increasing in utility offered at the margin in the metro area and hence at the margin for all localities in the metro area (given equalized utility for migrants across the metro area). Utility needs to rise as the total number of migrants, $\bar{L}$, to the metro area increases, with an inverse supply function of the form $f(\bar{L})$, $f' > 0$. Equating utility of the marginal unserviced migrant in our locality, $U = U(w(\bar{N} + \bar{L}), p, c, L_0, L, Z)$, to this inverse supply, we can solve for the locality housing price-level to get

$$p = p(L, c, L_0, Z; \overline{N} + \overline{L}). \quad (5)$$

Substituting for p from Eq. (5) into Eq. (4a), and rearranging we get

$$L = L(A, c_0, c, L_0, Z; \overline{N} + \overline{L}). \quad (6a)$$

Using metro area fixed effects (conditioning on $\bar{N} + \bar{L}$), we estimate a version of Eq. (6a) to show how the policy variable, L_0, as well as housing supply conditions, A, affect locality population, where we expect $\partial L / \partial L_0 > 0$.

So far we have implicitly assumed migrants are of one skill type: low-skilled. In actuality, we may also have in-migration of high-skilled individuals from outside the metro area as well as movements of existing high- and low-skilled residents across localities. While we estimate overall locality household growth equations based on Eq. (6a), we also separately estimate growth equations for high-skilled and low-skilled households and look at categories of migrants.

3.2. The choice of servicing for migrants

Suppose incumbent local residents of a locality are free to choose servicing for migrant households (i.e., incumbents choose L_0). Incumbents would be a set of wealthier long-term residents in each Brazilian locality in the 1980s who had sufficient control over the political process and selection of mayors so as to strongly influence exclusionary policies, including servicing of housing in migrant neighborhoods. We do not formally model this political process at the local level in this national non-democratic era, but instead just assume that "generic" incumbents are in control. In the empirics analyzing servicing choices, we control for characteristics of incumbents: locality median income and a measure of local political preferences. With democratization, however, localities came under strong pressure to make piped water universally available, and this presumption of discriminatory servicing is less plausible today.

For incumbents, we use a reduced-form specification of preferences,

$$V(y(\overline{N} + \overline{L}), \tau(L_0; \cdot), c_0, L_0, L, Z), \quad V_y \ , y_{\bar{L}}, \tau_{L_0}, V_{L_0} \geq 0; V_\tau, V_L < 0. \quad (7)$$

First, for incumbents as for migrants, we allow more migrants, $\bar{L}$, to the overall metro area to possibly increase real incomes of metro area residents, $y(\bar{N} + \bar{L})$, due to metro area agglomeration economies. Second, $V_{L0} \geq 0$, reflecting either altruism or positive externalities from better servicing of migrants in the locality. Third, $V_L < 0$, so that, while incumbent residents want more migrants to enter the metro area, they do not want them in their own locality. This could reflect local congestion considerations or prejudices against migrants. Finally, we assume without loss of generality that incumbents face the same costs of public water as migrants in serviced neighborhoods.

The second term in Eq. (7) is new and important. We assume incumbents are owner-occupiers of fully paid-off housing and therefore (7) must include the continuation value function for future periods (where the future is expected to be the same as today). Housing costs for incumbent residents incorporate their property taxes, which go towards financing locality infrastructure inclusive of the costs of water mains for those with newly-serviced housing. This is increasing in the number of serviced migrant units, L_0, and is also a function of past servicing in the locality. For example, if more houses were connected historically, new houses may be cheaper to service (e.g., just extend existing mains versus installing entirely new systems).

Incumbent residents choose service levels for migrants to maximize their utility. The first-order condition from (7) is

$$V_y y_{\bar{L}}(\partial\bar{L}/\partial L_0)+V_\tau \tau_{L_0}+V_{L_0}(\cdot)+V_L(\cdot)\partial L/\partial L_0=0, \tag{8}$$

which reflects trade-offs: better servicing in this locality may (1) increase metro area scale and agglomeration economies, (2) increase taxes of incumbents, (3) improve local environmental conditions (or appeal to altruism), and (4) result in increased congestion in the locality as more migrants are induced to enter.

A key issue is how to specify an estimating equation based on (8). If we assume "perfect competition" for residents among localities, then it is easy. The first term in (8) equals zero (i.e., $\partial\bar{L}/\partial L_0 = 0$) and the expression $\partial L/\partial L_0$ in the last term comes from Eq. (6a). We then can specify a form to the implicit function based on (6a), where

$$L_0=f(A, c_0, c, Z;\overline{N}+\overline{L}). \tag{9a}$$

In estimation of (9a), metro area fixed effects control for metro area scale; and the focus is on the role of items in Z, the socioeconomic characteristics of the locality in determining servicing. Of particular interest is how different localities set service levels for migrants according to whether they are higher-income versus lower-income or larger versus smaller, reflecting tax, externality, cost of water, and altruistic considerations.

An additional concern is that in metro areas with just a few localities, the assumption of "perfect competition" is less plausible and surely there may be strategic interactions. We briefly summarize the formulation with strategic interactions, which requires us to look at equilibrium in the whole metro area.

3.3. Equilibrium across localities within a metro area

In the metro area as a whole there are different localities, indexed 1, 2, 3, and so on. We have an Eq. (6a) for each locality, noting that, for metro area j, $\bar{L}_j = \Sigma_{i,i\in j} L_i$ where i indexes localities in the metro area. Given $\bar{L}_j = \Sigma_{i,i\in j} L_i$ Eq. (6a) for each locality within the metro area, in principle we can solve for a system of equations for each locality where:

$$L_{ij}=L_{ij}(\boldsymbol{L}_{0j},\boldsymbol{A}_j,\overline{N}_j,\overline{L}_j,\boldsymbol{Z}_j,\boldsymbol{c}_{0j},\boldsymbol{c}_j), \quad i=1,\ldots,n_j \ , \tag{10}$$

where n_j is the number of localities in metro area j; $\boldsymbol{L}_{0j}$, is the vector of service levels of localities in metro area j; $\boldsymbol{A}_j$ is the vector of land supply endowments for migrants in each locality in metro area j; $\boldsymbol{Z}_j$ is the vector of socioeconomic characteristics; and we have allowed serviced and unserviced water costs to vary across localities. The function is also indexed since, in principle, its form should depend on the number of localities, n_j, in the metro area. In (10), a locality's in-migration depends not just on its own servicing choices, but also on the choices of other localities in the metro area. We acknowledge that we are implicitly assuming (10) has a solution, and we are ignoring the possibility that such a solution is not unique. Existence and uniqueness would require a set of restrictions on functional forms and/or regions of parameter space.

Given Eq. (10), with strategic interactions, we return to Eq. (8) where we now have two new considerations as applied to any locality *i*. Now in the first term of (8), $\partial \bar{L}/\partial L_{0i} = \Sigma_{i,i\in j} \partial L_{ij}(\cdot)/\partial L_{0i}$, where the right-hand side is based on (10). Similarly, in the last term of (8), now $\partial L_i/\partial L_{0i}$ is based on (10). Eq. (8) utilizes the Nash assumption that a locality chooses its service level while holding other localities' choices fixed. For strategic interactions, we then rewrite (9a) as

$$L_{0i}=\tilde{f}(\boldsymbol{L}_{0,-ij},\boldsymbol{A}_j,\overline{N}_j,\overline{L}_j,\boldsymbol{Z}_j,\boldsymbol{c}_{0j},\boldsymbol{c}_j), \quad i=1,\ldots,n_j \ . \tag{9b}$$

Now choices of own servicing depend on service levels chosen in other localities, the strategic interaction part, as well as characteristics of other localities in the metro area. One can of course solve the system in (9b) to obtain a reduced-form version which subsumes strategic interactions, so that

$$L_{0i}=\widehat{f}(\boldsymbol{A}_j,\overline{N}_j,\overline{L}_j,\boldsymbol{Z}_j,\boldsymbol{c}_{0j},\boldsymbol{c}_j), \quad i=1,\ldots,n_j \ . \tag{9c}$$

In the empirical section, we focus on the estimation of Eqs. (9a) and (9b), depending on the assumptions we are prepared to make about the degree of competition. We also introduce other considerations, such as differing tastes across localities among incumbent residents concerned with inequitable provision of public services.

4. Effect of service provision on locality growth and composition

What is the effect of servicing decisions on migration? We explore this question with a specification that is based on a linearized version of Eq. (6a) with metro area fixed effects. Since there is both in and out-migration from localities, we look at growth in the number of households as the dependent variable, which reflects net as opposed to gross entry of residents. We also consider compositional effects and separately look at growth in numbers of low and high-education households. In robustness checks, we examine results for how servicing affects entry of rural and low-education migrants into a locality, capturing gross as opposed to net migration.

A key issue is how to characterize servicing. We look at small houses likely to be occupied by migrants and consider the extent to which they are serviced. While we try this as absolute counts of small houses serviced, as suggested by the model, to put it as a policy variable with a comparable index across localities, we specify the policy variable to be the share of small houses serviced in a locality, $b \equiv L_0/L$. Using an index proxies for a locality's *effort* at servicing migrant areas. A large locality may have absolutely many more unserviced small houses compared to a small one but may offer a much higher servicing rate; the index directly deals with this scaling issue. In the reduced-form Eq. (6a) from the model, we divide both sides by L to the power that L_0 is raised to (in the local log-linearized version) and rearrange, so that we have

$$L=g(A,c_0,c,Z;\overline{N}+\overline{L}), \quad b \equiv L_0/L. \tag{6b}$$

Although the policy variable at the margin was specified as the total number of serviced units in the theory section, in a well-behaved model and in (6b), it can just as readily be

NIH-PA Author Manuscript
NIH-PA Author Manuscript

specified as the share of serviced units, b. This policy variable, b, is increasing in L_0, and a decision about L_0 has a correspondence to a decision about b. Throughout the empirics, we rely on b as the policy variable describing servicing, although we experiment with alternatives.

The basic estimating equation is

$$\ln(L_{i,t}) - \ln(L_{i,t-1}) = \beta A_{i,t-1} + \gamma b_{i,t-1} + \varphi Z_{i,t-1} + \upsilon_j + \varepsilon_{i,t}. \quad (11)$$

We look at locality growth in the number of households between 1991 and 2000 as a function of locality characteristics in 1991. These include the rate of servicing of small houses, $b_{i,t-1}$, and a set of covariates, $A_{i,t-1}$, which describe housing supply and other conditions in the locality. Controls such as total migrants to the metro area, regional servicing standards, and metro area wages are swept into the metro area fixed effect, υ_j.

Another issue is the error structure. The urban growth literature (e.g., Glaeser et al., 1995) often takes the stance that (1) covariates are pre-determined and not affected by contemporaneous shocks that might induce growth, and (2) by looking at a growth equation, we have already differenced out time-invariant variables that affect long-run size. As such, in the literature, one standard approach is to rely on OLS estimation of cross-sectional growth equations. However, it seems likely that there are omitted variables affecting growth that persist over time, so that the $\varepsilon_{i,t-1}$, which affected past growth and the evolution of the pre-determined covariates, may be correlated with $\varepsilon_{i,t}$.

Of greatest concern is the possibility that the policy variable, $b_{i,t-1}$, is correlated with the error term, that is $E(b_{i,t-1}, \varepsilon_{i,t}) \neq 0$,causing the OLS estimates to be biased. Servicing today may be affected by past locality servicing, given scale economies in public water provision. For example, positive growth shocks in the past may have induced lower servicing as incumbents tried to restrict growth (or the locality faced capacity back-logs), making current expansion of servicing more expensive. Thus, low servicing in 1991 may reflect unmeasured good growth conditions from 1980–1991 for the locality, and such growth conditions may persist into the 1990s. That is, high past growth is negatively correlated with current supply of public water connections. Such influences will bias the OLS coefficients downward, understating the positive effects of good servicing on encouraging in-migration. The same issue relates to housing supply conditions: good unobservables driving locality growth in the past influence housing supply conditions today. Given this bias with OLS, we need to instrument for both the service variable and housing supply conditions in the locality.

4.1. Instruments

Since we have metro area fixed effects, instrumenting is challenging given the need to predict within metro area variation in covariates. We are focusing on locality growth between 1991 and 2000 and are concerned with past local economic shocks that affected servicing choices in the 1980s and persist into the 1990s. One set of instruments for contemporaneous locality characteristics will be 1970 metro area and locality characteristics discussed below. To be a valid strategy, use of such instruments requires that shocks which affected 1970 localities do not persist into 1991; and we will discuss why we think this is the case in the context of Brazilian development. A second set of instruments for contemporaneous service offerings will be climatic and geologic instruments, to which we now turn.

4.1.1. Extent of servicing—We use climatic and geologic instruments, following a similar strategy as Rosenthal and Strange (2008), to aid in predicting service levels. We start with the inferred history of water provision in the locality, which influences the costs of extending central water provision today. Ideally, we would like to instrument using a measure of how easy it would be for localities to provide public water connections. One such measure would be the degree to which incumbents are serviced by wells, which would imply low scale economies in central water provision. With wells, incumbents have an almost zero marginal cost of water and thus have limited demand for central provision for themselves. We unfortunately do not observe wells in the data, however we know the geology and weather variables which would have driven the efficacy of private water supply through wells. The idea is that localities with historically better conditions for wells will offer fewer central water connections today. However, given their high installation costs, wells are not an option for any low-income unserviced households in 1991, so those without a central connection must obtain water by hauling from stand-pipes or purchasing it privately.

What geologic conditions determine the efficacy of wells historically? First, if underlying sediments and rocks in a locality are more porous, they retain more water and wells are a more viable alternative to a public water connection (Food and Agriculture Organization, 2006). Second, the rate and variability of underground water replenishment will affect the viability of wells. As a measure of the potential for ground water replenishment, we use data on insolation, which is the amount and intensity of sunlight reaching the earth's surface. High insolation is associated with less rainfall and cloud cover and with more evaporation, impeding ground water replenishment. A high variance over the year in insolation means there are more intense periods of water replenishment without evaporation (Trimble and Ward, 2003). While insolation does vary across localities in a metro area, the fraction of the locality area having porous rocks and sediment varies much more.

Our key instruments for small houses served with a central water connection are the share of the locality area composed of porous geology (reducing the need for servicing, historically), this share porous geology interacted with mean insolation (increasing the need for servicing), and this share porous geology interacted with the standard deviation of insolation (reducing the need for servicing). Column 1 in Table B2 contains the first-stage regression showing that these instruments are strongly correlated with servicing, with the expected effects. While it is impossible to prove that these instruments only affect growth in the number of households through the channel of water supply, we feel we have been careful in our choice of instruments. We have focused on interactions of insolation with geology variables in case migration to a locality might be independently driven by either weather (a potential amenity) or geology (influencing construction costs). And our geology variable, measuring how porous are underlying rocks and sediments, is at a depth where wells are typically dug, which unlike surface soil measures, should not be expected to independently influence or be influenced by the economic landscape and growth of localities.[3] Thus we believe that instrumenting for servicing using geology and weather variables meets the exclusion restriction: these instrumental variables are correlated with public servicing levels but should have no direct, independent effect on current household growth rates.

4.1.2. Housing supply conditions—The second set of instruments are for variables relating to housing supply conditions and potentially other characteristics of the locality that might influence migrants' well-being. Controlling for total land area, these characteristics

[3]Since we use geology and not surface soil measures, this should alleviate some concern that our instruments might be correlated with factors such as agricultural potential, which might directly influence current growth. Also, while underlying geology might influence construction costs of large buildings with deep foundations, as in the instrumentation strategy in Rosenthal and Strange (2008), it is less of an issue in construction costs of single-family houses, and therefore should not directly affect the growth of households in our case.

include the number of urban households, average education (influencing the demand for space, as well as potentially being an amenity), and the share of households that are rural in the locality (influencing the potential supply of higher density urban housing). We instrument for these variables using 1970 metro area and locality characteristics. We now discuss why we think this is a valid strategy.

For validity, we require unobservables which affected metro area and locality growth in the past to be different from unobservables affecting growth today, so error drawings from 1970 need to be uncorrelated with drawings in the 1990s. Use of historical conditions as instruments introduces a tension between going further back in time to break the persistence in relevant unobservables and weakening the strength of the instruments. And the bias is clear: if past shocks (over and above metro area fixed effects) persist today so our historical instruments are not completely valid, we would still be understating the effects of covariates. In our case, we think the specific timing of development in Brazil makes 1970 a reasonable choice as a date for historical instruments. As discussed in Section 1, Brazil's initial rapid urbanization and spatial concentration of industry that occurs after World War II and extends into the 1970s is somewhat separate from today's more modern economy. By 1991, Brazil is 75% urbanized. The drivers of metro area growth have changed with the development of new export markets and new agricultural crops for export, as well as the move from heavy industry based on state capitalism to lighter industry based on the private manufacture of consumer products (Da Mata et al., 2005). Transport innovations and infrastructure investments have led to the decentralization of economic activity within metro areas, changing the economic activities of localities. These changes in the drivers of locality growth from before the 1970s to the 1990s means that historic locality characteristics should have no independent effect on growth between 1991 and 2000 except through their relationship with 1991 characteristics.[4]

As Table B2 in Appendix B shows, our instruments are reasonably strong. Past economic conditions of localities affected housing and other irreversible investment decisions as well as locality population and educational composition in the past. Historic accumulations are relevant since any adjustments away from them in locality characteristics are slow. If, in the 1960s, a locality attracted low-education migrants who settled in dense neighborhoods in the locality, that influences current educational composition even if locality economic conditions have changed completely. The instruments are implicit in the partial model we specified, which has simple dynamics—the past characteristics (denoted by $\boldsymbol{Z}$) influence the present with a given incumbent population. A more detailed version of the model would have full dynamics and explicitly allow for costly adjustments in residential markets within and across localities, limited reversibility of types of housing, and initial differences in stocks of high versus low-income residents.

For instruments, we include the following variables either on their own or interacted with other instruments: (1) the share of votes in the locality for anti-military parties in the 1982 national legislative elections, which reflects historical preferences of residents and may influence current policies governing housing supply; (2) access of a locality to Sao Paulo markets, which played a critical role historically, before the development of modern transnational transportation systems even though today it has little impact on growth;[5] (3) the illiteracy rate among the adult population in the locality and in the rest of the metro area in 1970, which influences, through accumulation, the average educational attainment today;

[4]In our work, it was clear that instruments from 1970 gave much better specification test results for the 1991–2000 period, compared to the 1980–1991 period, one reason why we focus on the 1991-2000 time period for population growth equations.

[5]We experimented with replacing distance to Sao Paulo with latitude. Results are very similar, but specification tests favored the original set of instruments.

(4) the manufacturing-to-service employment ratio in the rest of the metro area in 1970, which helped metro area economic attainment at the time and influences local economic composition today; (5) the number of households in the rest of the metro area in 1970, which provides a historical size measure influencing metro size today; and (6) the share of households that were rural in the rest of the metro area in 1970 and would be a basis for urban growth and later size. Note the attempt to generally rely on characteristics of localities in the rest of the metro area—i.e., in localities other than the own locality—in order to mitigate problems of persistence of own locality unobservables. We also note that many of these covariates such as illiteracy or socioeconomic status of the locality population, in addition to geologic conditions, will have influenced the use of wells in one locality compared to another in the past. A full set of first-stage regressions is reported in Table B2 and specification test results on validity of instruments are reported in the main tables and in the text.

4.2. Effects of servicing on the growth of urban households

Table 3 contains the basic results of the effects of servicing on the growth in the number of urban households in a locality between 1991 and 2000. Columns 1 and 2 contain OLS estimates, with shorter and longer lists of covariates. Columns 3 and 4 report corresponding 2SLS estimates. Standard errors are clustered at the metro area level to allow for correlation in contemporaneous shocks across localities within the metro area (after already controlling for metro area fixed effects). We experimented with a shorter instrument list and with estimation by Limited Information Maximum Likelihood (LIML). These results are discussed below but not reported in the table; however, we provide these results in the supplementary materials, which are available online. We also experiment with different samples and specifications of the outcome and policy variables.

4.2.1. Servicing—In OLS estimation in columns 1 and 2, the coefficient on servicing is negative and significant, reflecting the anticipated bias. Localities subject to the strong growth shocks of the late 1980s have poor servicing, potentially because of capacity expansion problems and past strategic choices. Instrumental variables estimation takes this negative coefficient and reverses its sign, making it positive. This positive coefficient is large. For a point estimate of about 0.73 in either column 3 or 4, a one standard deviation (0.21) increase in servicing leads to an increase of 0.15, or almost one standard deviation, in the growth rate in the number of households (for which the mean is 0.40) during the decade. This is a basic result of the paper: poor servicing of small houses likely to be occupied by low-income and low-education migrants has strong negative locality growth effects. Withholding water supply to retard locality growth below what it would have been in the absence of such withholding is effective.

However, the coefficient is somewhat noisily estimated, always significant at the 10% level but not quite at the 5% level (noting error terms are robust to heteroskedasticity and clustered at the metro area level). There are two issues. First is a weak instruments problem. While the partial F-statistics in first-stage regressions are satisfactory (see Table B2), for service levels the partial F is only 13.82 and it is hard to reject weak instruments overall (e.g., the Kleibergen–Paap LM statistic has a p-value of 0.17, meaning we cannot reject underidentification). We are instrumenting for multiple interrelated variables under metro area fixed effects, limiting the overall power of the instruments. LIML estimators are more robust to weak instruments than 2SLS but are more sensitive to the length of the instrument list. We obtain a much larger estimate of the servicing coefficient significant at the 5% level when estimating with LIML using the original instrument list, but obtain a similar estimate to that in column 3 of Table 3 when estimating with LIML and a shortened instrument list.[6]

The second issue has to do with the dependent variable: the total growth of households in a city. In the robustness checks, we discuss a variety of alternative measures, but the key issue is that exclusionary policies are aimed at lower-skilled migrants, not all migrants. In Section 4.3, we look at low-skilled households separately and obtain much sharper results.

4.2.2. Housing supply conditions—In Table 3, the basic controls on housing supply are land area and the number of households—controls for density of overall development. Again, the biases in moving from OLS to 2SLS estimation are what we expect. Having a high number of households is associated with recent strong local shocks and on-going growth, understating the negative effect of crowding on future housing supply conditions. So, for the number of households (increasing crowding), an OLS coefficient in column 1 of −0.025 becomes −0.112 under 2SLS estimation; and for land (alleviating crowding), the OLS coefficient of 0.006 becomes 0.059 under 2SLS. Additional variables in columns 2 and 4 which are education (reducing land supply but improving amenities) or share of households that are rural in the locality (increasing land supply), have no significant effects when estimating with 2SLS.

4.2.3. Robustness checks on growth equations—We perform robustness checks starting with a placebo policy variable and then moving onto the choice of samples and to definitions of the policy variable and outcome variables for the growth equation. These are presented in the supplementary materials, available online, but we discuss them in detail here. For the placebo, we replace the servicing of small houses with water by the fraction of those houses with a radio or TV set, which should have no causal effect on growth (controlling for fixed effects for differences in metro area incomes). The OLS coefficient for this variable is positive and significant as might be expected. Although instruments are weak (the first-stage *F*-statistic for the fraction of small houses with a radio or TV set is 7.10), the 2SLS coefficient is zero, not just insignificant. The 2SLS coefficient (standard error) is −0.017 (0.317), compared to the OLS coefficient of 0.257 (0.094). The ability of small households to have a radio or TV set should not lead to higher growth in the number of households, since radio and TV broadcasts are neither excludable (unlike public water connections) nor are they necessarily highly valued by the poor (i.e., in World Bank (2001), Brazil's poor listed “leisure” as being of low-importance in their list of wants and needs, far below housing).

In terms of samples, we reestimate the model in Table 3, separately dropping (1) the 44 localities in which there are multiple municipalities where the dominant municipality has less than 85% of the locality urban population and (2) all localities in the two largest metro areas, Rio de Janeiro and Sao Paulo. For sample (1), the idea is that we may have conflicting policies in localities with effectively more than one municipality, which could have weakened our results; and, for (2), large metro areas with huge numbers of localities may exhibit different functional responses than smaller metro areas. In samples (1) and (2), dropping the potentially problematic localities does not noticeably change results, although the loss of sample weakens significance levels. The Table 3, column 3 coefficient (standard error) on the service variable moves from 0.729 (0.419) to 0.654 (0.377) and to 0.601 (0.400), respectively, for samples (1) and (2).

[6]Estimating column 3 of Table 3 using LIML instead of 2SLS, we obtain estimates for the coefficient (standard error) on the servicing variable of 1.251 (0.660). When estimating column 3 of Table 3 using LIML and a shortened instrument list, we obtain estimates of 0.853 (0.507). With the short instrument list, the Kleibergen– Paap LM p-value is 0.04, although the precision of the coefficient estimate for the service variables is not improved. The shortened instrument list is adult illiteracy in 1970, share of the locality population that voted against the military in the 1982 legislative elections, the manufacturing-to-service employment ratio in the rest of the metro area in 1970, this ratio interacted with the log of the distance of the metro area to Sao Paulo, the log of the number of urban households in the rest of the metro area in 1970, the share of the locality's geology that is porous, this share porous interacted with mean insolation, and this share porous interacted with the standard deviation of insolation.

For the policy variable, rather than using the share of small houses serviced, we use the log total count of small houses serviced, with its attendant indexing and scaling issues noted earlier. Here we get a positive but insignificant OLS coefficient (standard error) of 0.044 (0.037), which rises to 0.149 (0.099) under 2SLS, suggesting that a 1% increase in the number of serviced small houses leads to a 0.15% increase in the growth of the locality. This indicates the same direction of effects but with less precision and weaker first-stage results.

Finally, we experiment with other outcome variables. So far we have looked at total changes in households in a locality. There are two objections. One is that household numbers change not just because of migration, but as extended families break-up and as mortality changes in the incumbent population. The second objection is that the policy variable is not focused on all migrants but on low-skilled or low-education ones. We have no ready way of dealing with non-migration changes to household numbers. We do experiment, however, with in-migration of households, but that ignores out-migration, so it is looking at gross rather than net flows. We also only have information on gross in-migration for the previous five rather than the previous ten years (with the base period covariates, when looking at growth between 1991 and 2000, being from ten years ago). The type of migrant issue is analyzed in more detail in the next section, where we look at the increase in the number of low versus high-education households across localities.

Looking at gross migration flows, we experiment with two measures. First is the count of all household heads who were in a rural location in another locality five years ago and second is the count of all low-education household heads who lived in another locality five years ago. A low-education household head is one who has not completed primary school (still about one-third of our stock of households in 2000), which is the borderline for functioning literacy. Actual raw migrant counts (based on the long-form census before being adjusted for the sample frame) are fairly small especially in less populous localities, so there is considerable noise. For these two dependent variables, the direction of bias is the same as before in moving from OLS to 2SLS, with initial negative OLS coefficients becoming positive. With the outcome variable defined as the log count of rural migrant household heads in the last five years, the 2SLS coefficient (standard error) on the share of small houses serviced is 2.358 (1.436), and when the outcome variable is the log count of low-education household heads who lived elsewhere five years ago, it is 1.918 (1.454). While servicing of small houses leads to higher growth in the number of households in a locality, it appears to lead to particularly strong in-migration of rural and low-education households (although these are noisily estimated). This takes us to our next set of results, which is looking at the effect of servicing small houses on the social composition of localities.

4.3. Composition effects

So far we have looked at how withholding service leads to a decline in the growth in the number of households overall. Because it is aimed at small houses, withholding service acts more specifically to discourage in-migration of low-education, low-income, or rural households. Looking at the net growth of low-education households involves a more precise or targeted group. The effect of withholding service to small houses on high-education households may be very different, since its effects really operate as an externality. Do high-skilled migrants moving into better neighborhoods of a locality suffer from poor servicing in nearby low-skilled migrant neighborhoods? For these two groups of households, low-education (incomplete primary school) and high-education (at least primary school completed), we look separately at the effect of servicing small houses on their growth rates.

The basic specifications are in Table 4. For housing supply, we control for base period overall density and household count of the relevant group whose growth we are investigating. Columns 1 and 2 deal with low-education household growth, using both OLS and 2SLS estimation, and columns 3 and 4 deal with high-education household growth, again using OLS and 2SLS estimation.

For the effect of servicing small houses on the growth in the number of low-education and high-education households, OLS coefficients on the servicing of small houses are negative, as expected and as was the case in the estimation of the overall growth equation. With 2SLS, the coefficient on servicing for the growth of low-education households becomes positive, very large, and significant. It is critical to our thinking that this servicing effect, when applied to a more targeted population, is statistically very strong and significant for the targeted population. A one standard deviation increase in servicing (0.21) increases the growth of low-education households by 0.15 under 2SLS estimation.

For high-education households between 1991 and 2000, under 2SLS estimation, the coefficient for servicing of small houses is also large, positive, and almost the same in magnitude as for low-education households. The estimate is much noisier, however, and not significant. Again this makes sense. Servicing of small houses gives some indication of negative externalities that might be faced by high-education households from poor servicing in low-skilled migrant neighborhoods, but it seems likely that the force of this externality would play out very differently in different localities, depending on precise geography and urban layout.

To further pursue the notion of externalities and differential effects of poor servicing of small houses on the growth of high-education households, we estimate a “within” ratio model where the ratio is the growth rate of low-education relative to high-education households.[7] Here the coefficient on servicing of small houses is negative but not statistically different from zero (the coefficient is −0.389, with standard error 0.346). These results, along with the results in columns 2 and 4 of Table 4, hint that poor servicing of small houses has adverse effects on the growth of all education households, suggesting a negative externality for high-education households. But most critically for low-education households, poor servicing is a significant deterrent of growth in their numbers.

5. Determinants of locality infrastructure servicing

In this section we examine how localities choose service levels, focusing on the implementation of the various forms of Eq. (9). For the base case, we start with a version without a test for strategic interactions, which assumes localities behave as “perfect competitors”, as in Eq. (9a). The equation $L_0 = f(A, c_0, c; Z; \bar{N} + \bar{L})$ in (9a) can be estimated for each locality, using metro area fixed effects to control for $\bar{N} + \bar{L}$.[8] In practice, the effects we are looking for— the role of locality size and income—in determining locality service levels will be fairly invariant to specification of Eq. (9). After discussing the base case results, we check for robustness, explore counterfactuals, and proceed to test for strategic interactions.

[7]This ratio represents a form of differencing and has the advantage of removing location observables and unobservables whose effects are common to both low-education and high-education groups. Letting N^k, $k \in \{L, H\}$, be the number of households that are low-education (L) or high-education (H), b be the share of small houses connected to water, and θ be any differential in slope coefficient between low and high-education households' response to the share of small houses serviced, the estimating equation is then:

$$\left[\ln\left(N_{i,t}^{L}\right) - \ln\left(N_{i,t-1}^{L}\right)\right] - \left[\ln\left(N_{i,t}^{H}\right) - \ln\left(N_{i,t-1}^{H}\right)\right] = \beta_0 \ln\left(N_{i,t-1}^{L}/N_{i,t-1}^{H}\right) + \beta_1 \ln density_{i,t-1} + \theta \ln\left(b_{i,t-1}\right) + \varepsilon_{i,t}.$$

[8]As such, one could also loosely interpret these as estimations of the reduced form equation in (9c), which subsumes strategic interactions. However, that interpretation would be for a specification which ignores the characteristics of other localities in the rest of the metro area, which do appear in (9c).

5.1. Base case results

We examine the determination of service levels for small houses in 1991, just after democratization, presuming they reflect policy decisions made in the 1980s under dictatorship. We assume locality elites in the 1980s have the ability to manipulate servicing of neighborhoods to encourage or discourage in-migration, but we allow differences in locality median income and political preferences to affect choices. We use locality and metro area attributes of 1980 as covariates in explaining 1991 service provision to ameliorate the problem of simultaneity bias. As in Section 3, in empirical implementation we replace the count of small houses serviced, L_0, with the share of small houses serviced, $b = L_0/L$, to address index and scaling issues across localities.

The base case without strategic interactions is in Table 5, columns 1 and 2. The short list of locality characteristics (the Z variables) includes median household income, number of urban households, and the interaction between the two, with variables in logarithmic form, and metro area fixed effects. For these key covariates, we believe service provision for other houses in a locality (the externality) is a normal good whose levels will rise with median locality income, although the degree of increase may diminish at higher income ranges. Second, the traditional urban literature hypothesizes that there are scale economies in public service provision, particularly in infrastructure provision which would lead more populous localities to provide services more cheaply, although any scale effects may diminish beyond a relatively small size (Hirsch, 1970). Apart from diminishing income and scale effects, we also anticipate that larger, richer localities may have more incentives to deflect migrants. So while they value greater servicing for small houses, captured by b, they may have a stronger aversion to the increased congestion which higher b generates. Second, richer households may not want the children of low-income and low-education migrants in local schools. Lastly, there may be fiscal reasons for richer localities to want to deflect low-income migrants, such as the dilution of any property tax base. While we might expect positive income and scale effects (in terms of public service provision), we expect the interaction between these two to be strongly negative. The notion is that these effects are specific to the servicing of houses most likely to be occupied by low-income migrants, and hence by inference are concerned with deterring in-migration during a period when deterring in-migration by withholding services is possible. This idea will be reinforced by a look at two counterfactuals later: one looking at servicing after democratization and one looking at servicing of large houses.

Base case results are reported in Table 5. Column 1 gives the results with a short covariate list, which includes only locality median household income, number of urban households, the interaction of the two, and metro area fixed effects. Column 2 adds in controls for political preferences for more egalitarian policies, density, and the overall level of servicing in the locality in 1970. We focus on the column 2 results.

Since our discussion of income and scale effects is detailed, we start with a brief look at other covariates, which have expected effects. Preferences for more egalitarian policies, taken as the share of voting in the locality in favor of anti-military parties during the 1982 national legislative elections, is associated with higher servicing of small houses. Higher density, which entails stronger negative externalities from poor water and sanitation conditions and also reduces costs per household of laying water mains, is associated with higher servicing. And lastly, higher overall servicing in the locality in 1970, driven by geologic conditions and 1970 factors during a different political regime, raises 1991 servicing of small houses.

Now turning to the key income and size variables, as expected, provision rises with locality income and size, but the interaction of the two is negative. For income effects, at one standard deviation below mean size (at 8.25), a two standard deviation increase in income (0.78) increases servicing by 0.12, while at one standard deviation *above* mean size (10.25), the same increase in income leads to an increase in servicing that is 30% smaller (at 0.08). For scale effects, coefficients indicate that for localities at one standard deviation *below* mean income (at 9.13), a two standard deviation increase in size (2.82) has a tiny effect on servicing (0.013), while at one standard deviation above mean income (at 9.91), the same increase in size leads to a *decline* of servicing of 0.044 (from a mean of 0.77). Since we do not expect scale diseconomies in provision of services in richer localities and since we do not expect income effects to diminish just because a locality is bigger versus smaller, we think the negative interaction effects are evidence of strategic exclusion in richer, larger localities. This interpretation is bolstered by results in Tables 6 and 7.

First we explore the negative income-scale interaction in more detail in Table 6. We divide localities into separate size and income quintiles to allow for non-linear income and scale effects; and then we interact these quintiles, creating 24 cells relative to the base. In Table 6, we report results controlling for the locality characteristics and metro area fixed effects of the Table 5, column 2 specification. As Table 6 shows, both income and scale effects generally increase monotonically across quintiles. Scale effects are statistically weak. Point estimates of income cell coefficients are significant and rise monotonically; but the big significant effect is moving out of the lowest income category and after that any increases in effects are modest. Interactions are always negative, but generally are smaller and insignificant except in high income-size cells, especially cell 5. Table 6 shows the relevant income-size interactions from the highest income and size quintiles. These interaction effects are large, strongly diminishing the scale and income effects. Moving from the lowest income and size quintiles to the highest, ignoring interactions, raises the share of small houses serviced by 0.32; the interaction reduces that increase by about 0.20 for income cell 5 and size cells 4 and 5. Again we interpret this as evidence of strategic exclusion.

5.2. Robustness and counterfactuals to the base case

We reestimated the basic models in Table 5 dropping localities where the dominant municipality had less than 85% of the urban locality population. We also reestimated the basic models dropping all localities in the two largest metro areas, Rio de Janeiro and Sao Paulo. In both cases, there is almost no change in the coefficients, although we lose statistical power because of smaller sample sizes.

For counterfactuals, if we are correct about the nature of incentives to deflect low-income migrants, then we should not observe the relationships in Table 5 in two “counterfactual” situations. First, we would not expect the same exclusionary motivations to operate in the servicing of large houses, which do not cater to low-income migrants, if we presume that localities have less or no desire to deflect the rich. Second, with democratization, the national government embarked on a wide-spread policy to upgrade urban slums and their servicing. Voting also became mandatory, making it more difficult for the elite of a locality to control local politics and decision-making. Thus, by 2000, after a decade of democracy and more representative local governance, we should see less or no evidence of localities withholding servicing to small houses.

As columns 1 and 2 of Table 7 show, there are no income and scale interactions in the servicing of large houses in 1991 and also none for small houses in 2000. For the quintile formulation corresponding to Table 6, income and size effects and their negative interactions are insignificant when looking at both large houses in 1991 and at small houses in 2000. Thus, the exclusionary effects for servicing of migrant housing by existing residents that we found in Tables 5 and 6 disappear in these two counterfactual situations.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

5.3. Strategic interactions

We have presented results in Tables 5–7 assuming localities within a metro area decide on a policy of servicing small houses based only on their own characteristics. This is the “perfect competition” assumption. In reality, localities within a metro area may interact strategically with one another, responding to each others' policies toward migration. In this section, we test for whether localities interact strategically in servicing small houses.

There are two main ways to test for strategic interactions, one being the local public finance approach and the other being the industrial organization (IO) approach. The IO literature has moved away from directly estimating strategic interactions to estimating structural parameters of the relevant cost and demand equations, from which to calculate strategic interactions. With the type of model we have outlined, there are some issues in following the IO approach. In the IO context, there is usually a single market with simple cost functions and a market demand function. In our context, there would be many functions to specify and identify structurally: a supply of migrants mediated through housing markets and public service conditions, and then preferences of incumbent residents for migrants and their service levels, taking externalities into account. There are also many markets (metro areas) with the number of participants (localities) varying in each one. Given this complexity, we have chosen to follow the local public finance literature and test for strategic interactions directly (as in Besley and Case (1995) and Case et al. (1993)) in a reduced form context.

The local public finance approach faces problems too, however. The first is that we can only identify strategic interactions locally, in the sense that the literature uses linearized specifications without interactive terms to estimate Eq. (9b). Second, across metro areas, there are different numbers of localities which generates different forms to the L_0 equation for each urban area, an issue ignored in the local public finance literature and one that does not exist in the IO literature with just one market. Of course, our formulations in Section 4.1 also face these issues. The third criticism has to do with the specific method used in the literature to estimate strategic interactions. This method has been criticized by Conley (2008) as producing “unstable” results.

A standard formulation to Eq. (9b), based on the public finance literature as reviewed by Brueckner (2000), for a locality 1 in metro area *j* in setting its servicing of small houses is

$$b_1 = \gamma \sum_{i \neq 1, i \in j} w_{1i} b_i + Z_1 \varphi + \varepsilon_1. \qquad (12)$$

In Eq. (12), $\Sigma_{i \neq 1, i \in j} w_{1i} b_i$ is a weighted sum of all other localities' choice of servicing in metro area *j*, where again, we have used the share of servicing, *b*, as opposed to the count of small houses serviced, L_0, as the policy variable. Note that as a simplification, Eq. (12) ignores characteristics of the rest of the localities in the metro area which appear in (9b), which is a reason to include metro area fixed effects in estimation. Writing Eq. (12) for all localities, and including metro area fixed effects, we have

$$\mathrm{b} = \gamma \mathbf{W} \mathbf{b} + \mathrm{Z} \varphi + \boldsymbol{\upsilon} + \varepsilon, \qquad (13)$$

where $\boldsymbol{\upsilon}$ is a vector of metro area fixed effects and **W** is the weighting matrix, with a weight of zero for the own locality and zeros for all other localities not in the same metro area as the own locality. Weights for other localities in the same metro area are normalized to sum to 1.

Weights are typically chosen arbitrarily. We report results using a traditional approach: equal weights for all localities within the metro area. But we prefer an approach based on the model. Eq. (9a) contains land supply, a measure of a locality's capacity to absorb new migrant households. We experiment with weighting by a land supply measure so that weights on other localities' servicing variables, b_i for $i \neq 1$ and $i \in j$, increases with land supply, A_i.[9] The intuition for land-supply weights is that we believe localities are in greater competition, and therefore place more weight, on other localities that have more available land with which to accommodate new migrant households. We use the inverse of historical population density for each locality as a weight indicating greater availability of land supply.

Estimating Eq. (13) by OLS would yield biased estimates. By construction, since ε_1 influences b_1, and since b_1 affects other localities' choices of their b_i, $\Sigma_{i \neq 1, i \in j} w_{1i} b_i$ is correlated with ε_1, introducing bias. As an example, if there is a metro area fixed effect as part of the error structure, the estimate of γ is biased downwards (becoming negative in the extreme of two localities per metro area). There are different solutions in the literature for how to estimate Eq. (13) (e.g., Fredriksson and Millimet, 2002); we use the most common approach, which fits our specification with metro area fixed effects. The idea is to move all **b** terms to the left-hand side of (13) to eliminate this correlation between the error terms and the choice variable. Doing this yields

$$\mathbf{b} = (\mathbf{I} - \gamma \mathbf{W})^{-1} \mathbf{Z} \varphi + (\mathbf{I} - \gamma \mathbf{W})^{-1} \boldsymbol{\upsilon} + (\mathbf{I} - \gamma \mathbf{W})^{-1} \varepsilon. \quad (14)$$

However, Eq. (14) ignores spatial correlation of the error terms within metro areas (over and above the fixed effect), which almost certainly arises. Here, for example, unobserved, correlated geographic factors across localities within a metro area might affect public infrastructure choices. Thus, following the literature, we assume an error structure of the form $\boldsymbol{\varepsilon} = \boldsymbol{\Psi} \mathbf{M} \boldsymbol{\varepsilon} + \boldsymbol{\xi}$, where $\xi \sim N(0, \sigma^2 \mathbf{I})$ and **M** is a matrix of spatial weights. For this specification, the estimating model becomes

$$\mathrm{b} = (\mathrm{I} - \gamma \mathrm{W})^{-1} \mathrm{Z} \varphi + (\mathrm{I} - \gamma \mathrm{W})^{-1} \boldsymbol{\upsilon} + (\mathrm{I} - \gamma \mathrm{W})^{-1} (\mathrm{I} - \boldsymbol{\Psi} \mathrm{M})^{-1} \xi. \quad (15)$$

For the **M** matrix, we use weights calculated from the inverse distance between pairs of localities in the metro area (normalized to sum to one), which gives greater weight to neighboring localities with similar geography. In estimation of (15), for **Z** and **W**, we use lagged covariates to deal with issues of contemporaneous correlation between errors and covariates; however, policy choices, **b**, are contemporaneous. Estimation of (15) uses a block structure since each metro area has different dimensions to the matrices and vectors (according to the number of localities).

We examined the existence of strategic interactions for the columns 1 and 2 formulations of Table 5. Results are almost identical for both formulations; and, in Table 8, we report those for the column 2 specification only, which includes the longer list of locality controls.

In Table 8, the basic results are in columns 1 and 2. The coefficient of interest is the estimate of γ from Eq. (15), which is the direct test of the existence of strategic interactions. This term is strongly negative and highly significant in column 1, using inverse density weights

[9]To see how land-supply weights arise from the model, assume the following functional forms: preferences of migrants to locality i are $x + g^{\gamma} + b_i^{\delta}$,, where x is a general consumption good and g is water; supply of migrants to the urban area is denoted by $\bar{L}$; land supply is $p_i = L_i A_t^{-1}$:; and preferences of existing residents of locality i are denoted by $C_i + b_i^{\delta} - Z_i L_i^{\varphi}$.. In a two locality urban area, $i \in (1, 2)$, these specifications yield the following reaction function: $b_1^{\delta} = \frac{A_2}{(1 + A_2)} b_2^{\delta} + B_1 Z_1^{1/(1-\varphi)} + C_1$, where C_i and B_i are locality-specific constants and other variables are as defined in Section 3. Note the weighting of b_2 by land supply, A_2, in locality 2.

for **W**, and in column 2, using equal weights for **W**. The interpretation of a negative coefficient is that an increase in servicing in one locality leads other localities to withhold their servicing. While this increase elsewhere helps the own locality by attracting more migrants to the metro area, enhancing overall scale economies in metro area labor markets, the own locality responds by withholding servicing in order to deflect incoming migrants to other localities. A more complex version of strategic interactions than allowed in formulations based on Eq. (15) would have incentives to deflect migrants vary by locality income recognizing more explicitly the degree of Tiebout-type stratification within metro areas. There is anecdotal evidence suggesting that the smaller, poorer, and more suburban localities of Brazil's major metro areas have been more willing to embark on slum upgrading projects than the larger and wealthier central city localities, perhaps suggesting some “coordination” within metro areas to upgrade further-out slums in place of upgrading central city ones (Inter-American Development Bank, 1995). With the caveat that there is some debate about how to properly estimate strategic interactions, we do find direct evidence of strategic interactions across localities, but not the positive interactions associated with a race to the bottom or top in servicing.

We now briefly discuss counterfactuals, robustness checks, and additional details on the estimation of strategic interactions. The estimates of strategic interactions based on the incorrect OLS formulation in Eq. (13) are positive without metro area fixed effects and negative with them. In column 3 of Table 8, we show that the negative strategic interactions we obtain using maximum likelihood estimation (MLE) are not just driven by the metro area fixed effects. Estimating Eq. (15) using MLE, and replacing metro area fixed effects with a long list of metro area controls yields a negative, although not statistically significant, estimate of γ. We also note that if we restrict the sample to metro areas with just 2–3 localities, the strategic interaction term is enhanced (the estimate of γ becomes −0.753, with standard error 0.075, using inverse density weights for **W**). We might expect this type of result: the strongest non-competitive behavior is likely to arise in cases when there are few localities in a metro area, making cooperation or collusion between localities more feasible. Finally, we note that in Table 8, columns 1 and 2, the spatial correlation of the error terms is negative. This may seem surprising, but recall we have metro area fixed effects, so this is within metro area spatial correlation, where relatively good conditions in one locality may mean relatively bad conditions in a neighboring one. If metro area fixed effects are removed, the spatial correlation in errors becomes strongly positive as we see in column 3.

In summary, we find evidence of strategic interactions between localities within a metro area when estimating Eq. (15) using MLE and metro area fixed effects. Strategic interactions are strongly negative and significant, suggesting that as one locality increases servicing, another locality responds by withholding servicing. These effects are robust to estimating using different weighting schemes, to estimating without metro area fixed effects but controlling for metro area characteristics, and are accentuated, as expected, when estimating using a sample of metro areas with only 2–3 localities.

6. Conclusions

The extensive literature on the exclusionary policies of local jurisdictions tends to focus on the exclusionary policies of localities in developed economies where informal housing markets do not exist. We have attempted, in our work, to examine exclusionary policies in a developing country framework where informal markets not only exist but are relatively prevalent, and thus provide an alternative to formal housing markets when localities attempt to enact exclusionary housing restrictions. In such a scenario, whatever legal housing restrictions are in place, migrants can still enter the informal housing sector. Beyond

NIH-PA Author Manuscript
NIH-PA Author Manuscript

enacting legal housing restrictions, localities can withhold public infrastructure services to the informal sector and thereby create a disincentive for migrants to enter.

We have examined these migration and exclusion dynamics using a sample of 327 localities in 54 metro areas in Brazil between 1980 and 2000. Migrants move to these metro areas for employment and choose the localities in which they will live. However, certain localities may seek to limit in-migration, especially when migrants are low-income and low-education relative to incumbent residents. We focus on the provision of public water connections to small (1–3 room) houses in which low-income and low-education migrants are most likely to live. We estimate the effect of withholding public water connections to small houses on the growth in the number of households in the locality and find that not servicing small houses is effective at reducing growth, especially for the targeted population of low-education households.

We then estimate the determinants of public water provision to small houses. We find that richer localities provide more servicing (a wealth effect), larger localities provide more servicing (a scale effect), but being both rich and large is associated with reduced servicing of small houses. These results help explain the emergence of "superstar" localities in Brazil despite the existence of an informal housing market. By withholding service from the small (informal sector) houses where low-education migrants are most likely to live, "superstar" localities can limit entry of low-education and low-income migrants, slow their growth in number of households, and over time, experience disproportionate increases in their median household incomes as mostly high-education and high-skilled migrants enter the locality.

We also find evidence of negative strategic interactions among localities within a metro area in setting services, meaning that if other localities raise service levels attracting more migrants to the metro area, that leads the own locality to reduce service levels in order to deflect migrants away from the locality.

These findings of exclusionary and strategic withholding of services from small houses applies to the period when Brazil was a dictatorship and when it was feasible for elite residents of a locality to co-opt local government decision-making. We find no evidence of exclusionary withholding of services from large (more than 8 room) houses during dictatorship, as would be expected if localities have no desire to limit in-migration of high-education and high-income households. And after a decade of democracy and mandatory voting laws, we no longer find evidence of exclusionary withholding of services from small houses in Brazil.

This paper hopefully provides some insight on the historic proliferation of unserviced, informal sector housing across even large and wealthy Brazilian localities and offers relevant parallels for other developing countries that are currently urbanizing. The existence of unserviced, informal housing sectors is not only a product of insufficient capacity or wealth, but can also be an outcome of historic strategic and exclusionary behavior by localities in determining services.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Bernardo Bloom, Andrew Foster, Will Strange, three anonymous referees, and seminar participants at Carnegie Mellon, University of Quebec, Brown University, the Philadelphia Federal Reserve, World Bank, United Nations, Brazilian Institute for Studies on Labor and Society, LACEA, and NEUDC for helpful comments and suggestions. We are grateful to Cassio Magalhaes, Flora Maravalhas, and Adam Storeygard for excellent research assistance, to Lynn Carlson for assistance with GIS, and to Cristina Costa, Daniel Da Mata, and Hyoung Wang for providing data.

Appendix A. Rent hedonics

Hedonic regressions to determine shadow prices or consumer willingness-to-pay for attributes apply to specific markets. Each locality has its own housing market, so in principle rent regressions to obtain consumer willingness-to-pay for housing and neighborhood attributes should be estimated separately for each locality. We look at Sao Paulo and Rio de Janeiro municipalities in 1980; these are the central cities of these two metro areas. We do not have data on prices of owner occupied units, but we do have price data on rental units, most of which are single family housing (rather than apartments). In the hedonic equations, we control for a variety of basic house characteristics: number of bedrooms, number of other rooms, urban versus rural location in the municipality, six types of wall construction materials, seven types of floors, eight types of roofs, and whether the unit is a single family residence. We then control for a variety of servicing features. One identification issue in estimation is that there may be unobserved neighborhood attributes that are correlated with servicing or even with house attributes. To minimize this problem, we insert district-level fixed effects, where Sao Paulo municipality has 56 districts and Rio de Janeiro municipality has 24 districts. The most recent year for which we can do this is 1980; later years either do not have rent data or do not have district identifiers. Another identification issue is that markets even within municipalities may be segmented, so that higher and lower-income households may face different prices. That is, the observed willingness-to-pay in the data by the "marginal" consumer will differ by market segment since the marginal consumer will be different. We look at hedonic regressions for the overall rental market and for the market utilized by the poorest 20% of households in each municipality.

For services, we do a full examination of all types and forms of services for the two municipalities separately for all renters and for those renters who are among the poorest 20% in each municipality. For the overall sample of renters, we also show a simplified version where we use the typical summary measures: central water connection and "full service" (any electricity, central water connection, and central sewer connection). Table A1 shows the basic results.

In Table A1, the reported coefficients reflect the percent by which rents rise for a given service. From columns 1 and 4, it is clear that in both Sao Paulo and Rio de Janeiro, there is a high premium for having central water piped into the house relative to no piped water at all. The premium for a central water connection with indoor plumbing is substantially more than for well water with indoor plumbing (12.4% and 24.4% more in Sao Paulo and Rio de Janeiro, respectively), presumably reflecting the greater reliability and year-round supply of public water. From columns 3 and 6, it is evident that these premiums for a central water connection remain high even when we restrict the estimations to the poorest 20% of households. Electricity garners a very large premium, even more so if it is metered (legal), indicating both reliable supply and higher (amperage) effective service. Central sewer is much more valued than septic systems, especially in Rio de Janeiro. Septic systems used only by one home raise premiums modestly above having no service, presumably reflecting the failure of septic systems in these dense localities, and septic systems shared with multiple homes garner no premium compared with the alternative of emptying sewage directly into ditches and streams. Clearly, there could be neighborhood conditions that vary within districts of these central cities that are correlated with covariates, but the results are suggestive.

Table A1

Rent hedonics.

	Sao Paulo				Rio de Janeiro			
	(1) All	**(2) All**	**(3) Poorest 20%**	**Mean for All**	**(4) All**	**(5) All**	**(6) Poorest 20%**	**Mean for All**
Dependent variable: Ln(Rent), 1980								
Water, central conn., indoor plumbing	0.389*** (0.021)		0.352*** (0.027)	0.89	0.413*** (0.022)		0.353*** (0.026)	0.91
Water, central conn., no indoor plumbing	0.093*** (0.022)		0.086*** (0.027)	0.06	0.095*** (0.023)		0.118*** (0.027)	0.06
Water, well, indoor plumbing	0.265*** (0.023)		0.285*** (0.029)	0.02	0.169*** (0.036)		0.251*** (0.048)	0.01
Water, well, no indoor plumbing	0.038 (0.023)		0.031 (0.029)	0.02	−0.089*** (0.033)		0.014 (0.040)	0.01
Central sewer	0.127*** (0.006)		0.073*** (0.010)	0.63	0.202*** (0.011)		0.130*** (0.015)	0.89
Septic system, own house	0.077*** (0.006)		0.097*** (0.010)	0.13	0.087*** (0.024)		−0.022 (0.041)	0.01
Septic system, shared with others	−0.043*** (0.007)		−0.027** (0.010)	0.06	0.004 (0.036)		0.042 (0.044)	0.00
Electricity, metered	0.302*** (0.013)		0.253*** (0.021)	0.65	0.304*** (0.022)		0.354*** (0.035)	0.77
Electricity, unmetered	0.018 (0.013)		0.102*** (0.020)	0.34	−0.017 (0.022)		0.158*** (0.035)	0.21
Water from general system		0.182*** (0.007)		0.95		0.222*** (0.015)		0.96
Full service (water, sewer, electricity)		0.195*** (0.004)		0.60		0.362*** (0.007)		0.84
Controls: House Characteristics, District Fixed Effects	Yes	Yes	Yes		Yes	Yes	Yes	
Number of Households	201,001	201,001	46,585	201,001	109,166	109,166	25,149	109,166
Districts	56	56	56		24	24	24	
R^2	0.53	0.50	0.35		0.55	0.54	0.36	

Notes: Robust standard errors are reported in parentheses. Since the poorest 20% of households are disproportionately renters, they comprise more than 20% of the sample.

* Significant at 10% level.

** Significant at 5% level.

*** Significant at 1% level.

Appendix B. Summary statistics and first-stage regressions

Please refer to Tables B1 and B2.

Table B1

Summary statistics.

	Mean	Std. dev.	Min	Max
*For*Tables 3 and 4: N = *327, 54 metro areas*				
Growth rate of urban HHs, 1991–2000	0.397	0.172	−0.066	1.048
Growth rate of low educ. HHs, 1991–2000	0.536	0.196	−0.161	1.199
Growth rate of high ecu. HHs, 1991–2000	0.325	0.223	−0.255	1.294
Share small houses with water connection, 1991	0.743	0.213	0.022	1.000
ln # urban HHs, 1991	9.505	1.461	6.103	14.733
ln # low educ. HHs, 1991	8.510	1.336	5.428	13.321
ln # high educ. HHs, 1991	9.001	1.565	5.391	14.453
ln land area, 1991	5.976	1.358	2.797	12.226
ln density, 1991	5.111	1.731	0.218	9.416
Avg. education (years) in locality, 1991	4.712	1.100	2.059	8.840
Share HHs rural in locality, 1991	0.141	0.131	0.000	0.507
*For*Tables 5–8: N = *276, 50 metro areas*				
Share small houses with water connection, 1991	0.778	0.190	0.022	1.000
Share small houses with water connection, 2000	0.857	0.144	0.097	1.000
Share large houses with water connection, 1991	0.910	0.122	0.178	1.000
ln # urban HHs, 1991	9.253	1.407	6.297	14.523
ln # urban HHs, 1980	9.811	1.331	6.338	14.733
ln median HH income, 1980	9.517	0.388	8.412	10.440
ln median HH income, 1991	7.104	0.380	6.001	7.937
Share anti-military vote in 1982 national elections	0.585	0.152	0.029	0.941
ln density, 1980	5.104	1.639	0.636	9.360
ln density, 1991	5.466	1.643	0.846	9.416
Share all urban HHs with water, 1970	0.473	0.303	0.000	0.955
ln # urban HHs in metro area, 1980	12.188	1.569	8.916	14.886

Table B2

First-stage regressions.

	(1) Share small houses with water 1991	(2) ln # urban HHs 1991	(3) Avg. educ. in locality 1991	(4) Share HHs rural in locality 1991	(5) ln # low educ. HHs 1991	(6) ln # high educ. HHs 1991	(7) ln density 1991
ln land area 1991	−0.019 (0.011)	0.349*** (0.059)	0.079** (0.030)	0.021*** (0.008)			
Illiteracy rate in locality 1970	−0.007*** (0.001)	−0.074*** (0.008)	−0.077*** (0.007)	0.004*** (0.001)	−0.044*** (0.007)	−0.072*** (0.008)	−0.089*** (0.010)
Illiteracy rate in rest of metro area 1970	−0.025 (0.016)	−0.007 (0.132)	0.163 (0.120)	0.017 (0.012)	−0.108 (0.109)	−0.069 (0.123)	0.090 (0.168)

	(1) Share small houses with water 1991	(2) ln # urban HHs 1991	(3) Avg. educ. in locality 1991	(4) Share HHs rural in locality 1991	(5) ln # low educ. HHs 1991	(6) ln # high educ. HHs 1991	(7) ln density 1991
Share anti-military vote 1982 national elections	0.158 (0.099)	2.068*** (0.455)	0.046 (0.272)	−0.184*** (0.061)	1.947*** (0.374)	2.032*** (0.427)	2.040*** (0.615)
Manufacturing-service ratio, rest of metro 1970	−0.136 (0.293)	2.724** (1.101)	−0.284 (1.345)	−0.056 (0.080)	3.565*** (1.207)	3.329*** (1.233)	1.476 (1.063)
Manu-serv ratio rest of metro * ln dist to S.P. 1970	0.015 (0.054)	−0.446** (0.206)	0.136 (0.243)	0.015 (0.015)	−0.610** (0.231)	−0.538** (0.235)	−0.230 (0.197)
Illiteracy rate rest of metro * ln dist to S.P. 1970	0.002 (0.002)	−0.012 (0.017)	−0.029* (0.016)	−0.001 (0.002)	−0.001 (0.014)	−0.009 (0.016)	−0.017 (0.021)
ln # urban HHs rest of metro area 1970	−0.090** (0.037)	−0.946** (0.388)	−0.282** (0.136)	0.033 (0.029)	−1.554*** (0.362)	−1.702*** (0.378)	0.345 (0.348)
Share HHs rural in rest of metro area 1970	−0.042 (0.169)	1.719 (1.753)	2.329** (0.993)	−0.370** (0.154)	0.198 (1.447)	0.249 (1.535)	3.905* (2.082)
Share area porous geology	−0.493** (0.237)	−0.634 (1.437)	−0.712 (0.969)	−0.245 (0.209)	−0.552 (1.582)	−1.227 (1.473)	−0.244 (1.582)
Mean annual insolation	−0.099 (0.203)	0.151 (1.008)	0.326 (0.736)	−0.224 (0.164)	0.697 (1.108)	0.280 (1.193)	−0.633 (1.492)
Std dev of annual insolation	0.113 (0.458)	6.683 (5.512)	4.108*** (1.237)	−1.180** (0.534)	6.450 (4.878)	6.888 (4.781)	4.869 (6.728)
Share porous geology * mean insolation	0.336*** (0.108)	0.249 (0.710)	0.368 (0.536)	0.143 (0.117)	0.180 (0.752)	0.613 (0.716)	0.057 (0.867)
Share porous geology * std dev of insolation	−0.591*** (0.187)	−1.298 (1.191)	−1.045 (1.027)	0.066 (0.174)	−1.156 (1.373)	−1.686 (1.306)	−1.015 (1.617)
N [54 metro areas]	327	327	327	327	327	327	327
R^2	0.29	0.67	0.72	0.39	0.57	0.65	0.50
F-statistic	14.96	125.32	75.97	15.89	58.35	102.51	49.99

Notes: Robust standard errors, clustered at the metro area level, reported in parentheses.

* Significant at 10% level.

** Significant at 5% level.

*** Significant at 1% level.

Appendix C. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi: 10.1016/j.jue.2010.09.006.

References

Arretche, M. Water Supply and Sanitation. Ministerio das Relacoes Exteriores; Brasilia, Brazil: 2001.

Baer, W. The Brazilian Economy: Growth and Development. fifth. Praeger; Westport, CT: 2001.

Besley T, Case A. Incumbent behavior: vote seeking, tax setting, and yardstick competition. American Economic Review. 1995; 85:25–45.

Biderman, C. Regulation and Informal Settlements in Brazil: A Quasi-Experimental Approach. Lincoln Land Institute of Land Policy; Cambridge, MA: 2007.

Black D, Henderson JV. A theory of urban growth. Journal of Political Economy. 1999; 107:252–284.

Brueckner J. Welfare reform and the race to the bottom: theory and evidence. Southern Economic Journal. 2000; 66:505–525.

Cai, F. Floating Population: Urbanization with Chinese Characteristics. Institute of Population and Labor Economics; Beijing, China: 2006.

Case A, Rosen H, Hines J. Budget spillovers and fiscal policy interdependence: evidence from the states. Journal of Public Economics. 1993; 52:285–307.

Conley, T. Spatial econometrics. In: Durlauf, S.; Blume, L., editors. The New Palgrave Dictionary of Economics Online. second. Palgrave Macmillan; 2008. <http://www.dictionaryofeconomics.com/article?id=pde2008_S000450>

Da Mata D, Deichmann U, Henderson JV, Lall S, Wang H. Examining growth patterns of Brazilian localities. 2005 World Bank Policy Research Working Paper # 3724.

Da Mata D, Deichmann U, Henderson JV, Lall S, Wang H. Determinants of city growth in Brazil. Journal of Urban Economics. 2007; 62:252–272.

Dowall, D. Brazil's Urban Land and Housing Markets: How Well are They Working?. Lincoln Institute of Land Policy; Cambridge, MA: 2006.

Food and Agriculture Organization. World Reference Base for Soil Resources. Food and Agricultural Organization of the United Nations; Rome, Italy: 2006.

Fredriksson P, Millimet D. Strategic interaction and the determinants of environmental policy across U.S. states. Journal of Urban Economics. 2002; 51:101–122.

Gabaix X. Zipf's law for cities: an explanation. Quarterly Journal of Economics. 1999; 113:739–797.

Glaeser E, Scheinkman J, Shleifer A. Economic growth in a cross section of cities. Journal of Monetary Economics. 1995; 36:117–143.

Gyourko J, Mayer C, Sinai T. 2006 Superstar cities, NBER Working Paper # 12355.

Heller, L. Water and sanitation policies in Brazil: historical inequalities and institutional change. In: Castro, JE.; Heller, L., editors. Water and Sanitation Services: Public Policy and Management. Earthscan; London, UK: 2009. p. 321-337.

Hirsch, W. The Economics of State and Local Government. McGraw-Hill; New York, NY.: 1970.

Inter-American Development Bank. Brazil: Rio de Janeiro Urban Upgrading Program. Inter-American Development Bank; Washington, DC: 1995. Operation No. 898/OC-BR. Approved on November 1, 1995 (BR-0182)

IPEA, IBGE, UNICAMP. Configuração Atual e Tendências da Rede Urbana, Serie Configuração Atual e Tendências da Rede Urbana. Instituto de Pesquisa Econômica Aplicada, Instituto Brasileiro de Geografia e Estatistica, Universidade Estadual de Campinas; Brasilia, Brazil: 2002.

Rosenthal S, Strange W. The attenuation of human capital spillovers. Journal of Urban Economics. 2008; 64:373–389.

Scheper-Hughes, N. Death without Weeping: The Violence of Everyday Life in Brazil. University of California Press; Berkeley, CA: 1993.

Trimble, S.; Ward, A. Environmental Hydrology. second. CRC Press; Boca Raton, FL: 2003.

Wildasin D. Income redistribution in a common labor market. American Economic Review. 1991; 81:757–774.

World Bank. Brazil: Attacking Brazil's Poverty. Vol. 2. Washington, DC: 2001. Report No. 20475-BR World Bank

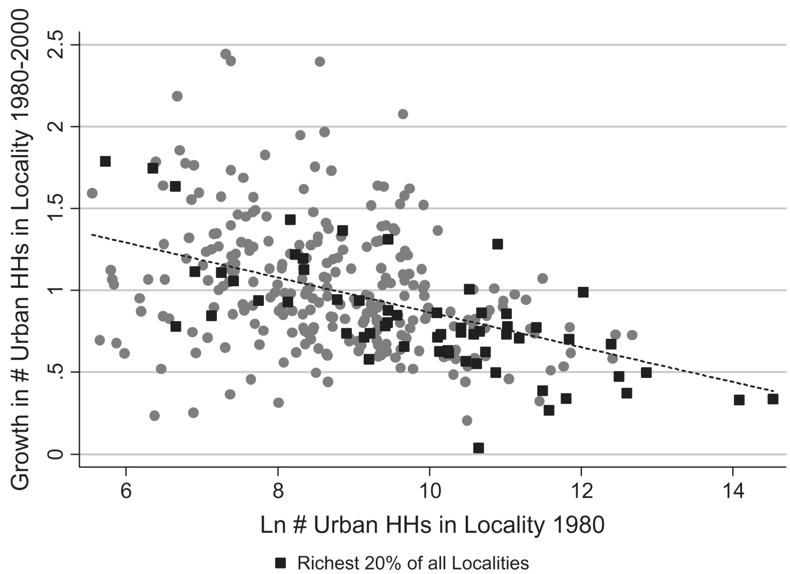

Fig. 1.
Growth in a locality's number of urban households, 1980–2000.
Notes: Slope coefficient (standard error) of -0.106 (0.013). R-squared=0.18.

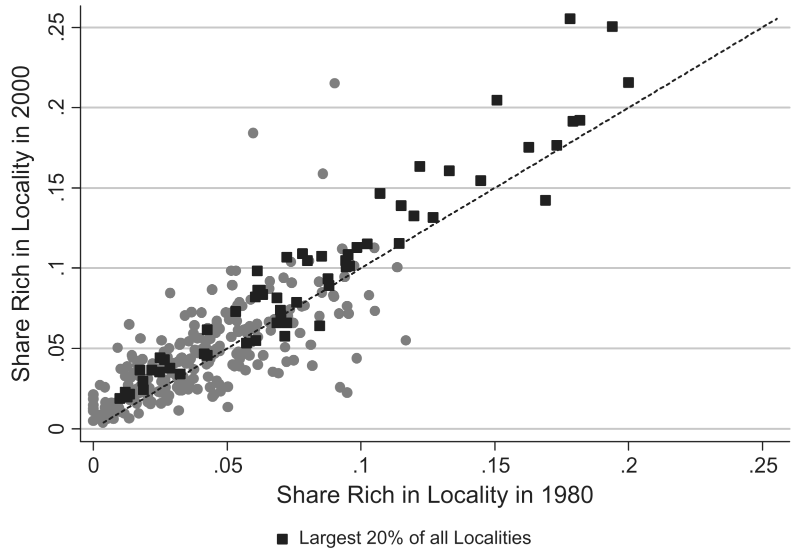

Fig. 2.

Share rich in a locality in 1980 versus 2000.

Notes: Regression coefficients (standard errors): ShareRich2000 = [0.013 (0.002)] + [0.802 (0.045)]×ShareRich1980 + [0.233 (0.041)]×ShareRich1980×DummyLargest20%. R-squared=0.77.

Table 1

Share of urban households with connection to public water service.

	Percent with central water connection (locality average)	Percent with central water connection (weighted by number of urban HHs in locality)	Number of localities with water	Total number of localities
Average for sample of 327 localities in 54 multi-locality metro areas, at least 50% urbanized by decade				
1970	50	58	156	185
1980	69	80	271	278
1991	83	91	326	327
2000	89	92	326	326
1991 Breakdown				
Own house but not land	74	82	326	327
Own house and land	85	91	326	327
Small houses	74	83	326	327
Large houses	89	95	326	327
Poorest 20% of Households	76	85	326	327
Richest 20% of Households	89	95	326	327
Average for localities in multi-locality metro areas, at least 50% urbanized by 1970				
1970	50	58	156	185
1980	75	82	183	185
1991	89	92	185	185
2000	92	93	185	185
Average for localities in multi-locality metro areas, less than 50% urbanized by 1970				
1970	34	47	94	142
1980	53	63	128	142
1991	76	78	141	142
2000	84	83	142	142

Notes: We consider localities that have less than 10% of houses serviced as having no service. Even though there may be in fact zero service, some households may have mistakenly declared in the census that they had service or service may be incipient. We weight the share of servicing in each locality by the number of urban households in the locality in each decade. When the weighted share of urban households serviced is greater than the unweighted share, this implies that the mean level of servicing is higher in larger localities.

Table 2

Public water servicing of houses by house size.

	1970	1980	1991	2000
Average for 327 cities in 54 multi-locality metro areas at least 50% urbanized by 1991				
Small houses	28	49	74	84
Large houses	61	78	89	91
Difference in servicing	33	29	15	7

Notes: Small houses represent approximately the bottom 15% of all houses based on total number of rooms by decade; large houses represent approximately the top 15%. Exact numbers are provided in the text.

Table 3

Growth rate of urban households, 1991–2000.

	(1) OLS	(2) OLS	(3) 2SLS	(4) 2SLS
Dependent variable: Growth rate of urban households, 1991–2000				
Share small houses with water 1991	−0.180* (0.091)	−0.162** (0.076)	0.729* (0.419)	0.739* (0.429)
ln # urban HHs 1991	−0.025*** (0.009)	0.014 (0.024)	−0.112*** (0.030)	−0.114* (0.062)
ln land area 1991	0.006 (0.010)	−0.012 (0.015)	0.059*** (0.019)	0.063* (0.034)
Avg. education in locality 1991		−0.040*** (0.015)		−0.005 (0.049)
Share HHs rural in locality 1991		0.262 (0.227)		−0.082 (0.494)
Metro area fixed effects	Yes	Yes	Yes	Yes
N [54 metro areas]	327	327	327	327
R^2 within metro areas	0.10	0.14		
Hansen–Sargan stat. *p*-value			0.94	0.87
Underidentification test *p*-value			0.17	0.12
Min. 1st stage partial F			13.82	13.82

Notes: Robust standard errors, clustered at the metro area level, reported in parentheses.

Excluded instruments are: locality adult illiteracy rate in 1970, adult illiteracy rate in the rest of the metro area in 1970, the share of locality votes for anti-military parties in the 1982 national legislative elections, the manufacturing-to-service employment ratio in the rest of the metro area in 1970, this manufacturing-to-service ratio interacted with the log of the distance of the locality to Sao Paulo, the adult illiteracy rate in the rest of the metro area in 1970 interacted with the log of the distance to Sao Paulo, the log number of households in the rest of the metro area in 1970, the share of households that are rural in the rest of the metro area in 1970, the share of a locality's land that is composed of porous geology, mean annual insolation in the locality, the standard deviation of average monthly insolation in the locality, the share of a locality's geology that is porous interacted with mean insolation, and the share porous geology interacted with the standard deviation of insolation.

* Significant at 10% level.

** Significant at 5% level.

*** Significant at 1% level.

Table 4

Growth rate of households by education level, 1991–2000.

	Low-education households		High-education households	
	(1) OLS	**(2) 2SLS**	**(3) OLS**	**(4) 2SLS**
Dependent variable: Growth rate of households by education level, 1991-2000 [ln(# low or high educ HHs)(t) − ln(# low or high educ HHs)(t − 1)]				
Share small houses with water 1991	−0.035 (0.098)	0.760*** (0.253)	−0.284*** (0.100)	0.715 (0.638)
ln density 1991	0.011 (0.013)	−0.029 (0.031)	−0.004 (0.012)	−0.099* (0.060)
ln # low educ. HHs 1991	−0.026*** (0.010)	−0.060*** (0.019)		
ln # high educ. HHs 1991			−0.018* (0.011)	−0.037 (0.028)
Metro area fixed effects	Yes	Yes	Yes	Yes
N [54 metro areas]	327	327	327	327
R^2 within metro areas	0.03		0.11	
Hansen–Sargan stat. *p*-value		0.48		0.76
Underidentification test *p*-value		0.28		0.25
Min. 1st stage partial *F*		13.99		13.99

Notes: Robust standard errors, clustered at the metro area level, reported in parentheses.

Excluded instruments are: locality adult illiteracy rate in 1970, adult illiteracy rate in the rest of the metro area in 1970, the share of locality votes for anti-military parties in the 1982 national legislative elections, the manufacturing-to-service employment ratio in the rest of the metro area in 1970, this manufacturing-to-service ratio interacted with the log of the distance of the locality to Sao Paulo, the adult illiteracy rate in the rest of the metro area in 1970 interacted with the log of the distance to Sao Paulo, the log number of households in the rest of the metro area in 1970, the share of households that are rural in the rest of the metro area in 1970, the share of a locality's land that is composed of porous geology, mean annual insolation in the locality, the standard deviation of average monthly insolation in the locality, the share of a locality's geology that is porous interacted with mean insolation, and the share porous geology interacted with the standard deviation of insolation.

* Significant at 10% level.

** Significant at 5% level.

*** Significant at 1% level.

Table 5

Locality determinants of service provision.

	(1)	(2)	(3)
Dependent variable: Share small houses with water connection in own locality, 1991			
ln median HH income 1980	0.477*** (0.134)	0.371*** (0.113)	0.518*** (0.147)
ln # urban HHs 1980	0.287** (0.134)	0.242** (0.105)	0.358** (0.154)
ln median HH income * ln # urban HHs 1980	−0.028** (0.014)	−0.026** (0.011)	−0.039** (0.016)
Share anti-military vote 1982 national elections		0.186** (0.091)	0.158 (0.106)
ln density 1980		0.032** (0.014)	0.045*** (0.014)
Share all urban HHs with water 1970		0.063 (0.039)	0.145*** (0.042)
ln # urban HHs in metro area 1980			−0.046** (0.021)
Metro area fixed effects	Yes	Yes	No
Extended list of metro area controls	N/A	N/A	Yes
N [50 metro areas]	276	276	276
R^2 [within metro areas]	[0.25]	[0.32]	0.48

Notes: Robust standard errors, clustered at the metro area level, reported in parentheses.

Extended metro area controls are: natural log of the distance from the metro area to Sao Paulo, natural log of the metro area land area, the illiteracy rate in the metro area, and the manufacturing-to-service employment ratio in the metro area, all for 1980.

* Significant at 10% level.

** Significant at 5% level.

*** Significant at 1% level.

Table 6

Income-size interactions in the setting of locality service levels.

	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Dependent variable: Share small houses with water connection in own locality, 1991				
ln median HH income 1980	0.137*** (0.048)	0.160*** (0.054)	0.165** (0.073)	0.213*** (0.068)
ln # urban HHs 1980	0.041 (0.061)	0.082 (0.056)	0.121* (0.064)	0.108 (0.071)
		Relevant Interactions		
4th income*4th size quintile		−0.133* (0.071)		
4th income*5th size quintile		−0.090 (0.089)		
5th income*4th size quintile		−0.228*** (0.070)		
5th income*5th size quintile		−0.190*** (0.067)		
Metro area fixed effects		Yes		
Locality controls from Table 5, column 2		Yes		
N [50 metro areas]		276		
R^2 within metro areas		0.37		

Notes: Robust standard errors, clustered at the metro area level, reported in parentheses.

Locality controls are: share of votes for the anti-military party in the 1982 national elections, natural log of population density in 1980, and share of all urban households with connection to water in 1970. The omitted categories are quintile 1 of income and size, and the omitted interaction is 1st income" 1st size quintile.

* Significant at 10% level.

** Significant at 5% level.

*** Significant at 1% level.

Table 7

Counterfactuals.

	(1) Service to large houses, 1991	(2) Service to small houses, 2000
Dependent variables: Share large houses with water connection in own locality, 1991; share small houses with water connection in own locality, 2000 (Covariates lagged by one period unless noted otherwise)		
ln median HH income	0.092 (0.093)	0.030 (0.138)
ln # urban HHs	0.055 (0.084)	−0.003 (0.093)
ln median HH income * ln # urban HHs	−0.005 (0.008)	−0.001 (0.012)
Share anti-military vote 1982 national elections	0.027 (0.070)	0.153 * (0.084)
ln density	0.024 *** (0.008)	0.035 *** (0.011)
Share all urban HHs with water 1970	0.045 (0.034)	0.098 ** (0.043)
Metro area fixed effects	Yes	Yes
N [50 metro areas]	276	276
R^2 within metro areas	0.29	0.24

Notes: Robust standard errors, clustered at the metro area level, reported in parentheses.

* Significant at 10% level.

** Significant at 5% level.

*** Significant at 1% level.

Table 8

Strategic interactions in servicing of small houses.

	(1) MLE Inverse Density Weights	(2) MLE Equal Weights	(3) MLE Inverse Density Weights
Dependent variable: Share small houses with water connection in own locality, 1991			
Weighted avg. other localities share of servicing of small houses	−0.590*** (0.081)	−0.590*** (0.133)	−0.184 (0.120)
ln median HH income 1980	0.252** (0.104)	0.245** (0.109)	0.413*** (0.114)
ln # urban HHs 1980	0.170* (0.102)	0.154 (0.107)	0.242** (0.111)
ln median HH income * ln # urban HHs 1980	−0.018* (0.011)	−0.017 (0.011)	−0.026** (0.012)
Spatial correlation of error terms	−0.290** (0.118)	−0.290* (0.159)	0.595*** (0.065)
Locality controls	Yes	Yes	Yes
Metro area controls	N/A	N/A	Yes
Metro area fixed effects	Yes	Yes	No
N [50 metro areas]	276	276	276
R^2	0.80	0.80	0.63

Notes: Standard errors reported in parentheses. Errors are corrected for spatial correlation within metro areas. Weights refer to the weighting of other localities' share of servicing of small houses in 1991 for localities within the same metro area. Locality controls are: share of votes for the anti-military party in the 1982 national elections, natural log of population density in 1980, and the share of all urban households with connection to water in 1970. Metro area controls are: natural log of the distance from the metro area to Sao Paulo, natural log of the metro land area, the illiteracy rate in the metro, and the manufacturing-to-service employment ratio in the metro area, all for 1980.

*Significant at 10% level.

**Significant at 5% level.

***Significant at 1% level.