Exclusionary Policies in Urban Development: Under-Servicing Migrant Households in Brazilian Cities*

Leo Feler
J. Vernon Henderson

September 22, 2009

Localities in developed countries often enact regulations to restrict supply and raise the price of housing, deterring low-income households from moving in. In developing countries, formal sector supply restrictions lead to informal housing sectors not governed by these restrictions. To deter low-income migrants, localities in developing countries withhold public services to the informal sector, where low-income migrants live, creating a disincentive for them to enter. Using a large sample of Brazilian localities, we examine migration and exclusion, focusing on the public provision of water to small houses in which migrants are likely to live. Withholding water connections reduces the growth rate of low-education households; but, perhaps because of negative externalities, it also reduces the growth rate of high-education households. In terms of service provision, in the pre-democracy era in Brazil, while richer localities provided more servicing to migrant housing (a wealth effect) and larger localities provided more servicing (a scale effect), being both rich and large was associated with reduced servicing, which we interpret as an intention to deter in-migration. Such effects are not found for non-migrant housing or during the democratic era.

Keywords: Urban Slums, Urban Growth, Urban Services, Exclusion.
JEL Classification: D7, H7, J6, O15, O54, R5.

Leo Feler
Department of Economics
Brown University
Box B
Providence, RI 02912
401-499-3447
leo_feler@brown.edu

J. Vernon Henderson
Department of Economics
Brown University
Box B
Providence, RI 02912
401-863-2886
vernon_henderson@brown.edu

* We thank Andrew Foster and seminar participants at Carnegie Mellon, Quebec, and Brown Universities, and at the Philadelphia Federal Reserve, World Bank, United Nations, Brazilian Institute for Studies on Labor and Society (IETS), Latin American and Caribbean Economic Association Conference, and Northeast Universities Development Consortium Conference 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.

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, maybe 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 populations in slums or informal housing sectors. For Brazil this is illustrated by the growth of its now infamous favelas and loteamentos. Today, these types of settlements are a growing phenomenon in South and South-East Asia, Sub-Saharan Africa, and in the urban villages of Beijing and other Chinese urban areas. Informal housing sectors are usually characterized by low housing quality and varying degrees of tenure insecurity. Perhaps more critically, they tend to be cut off from basic urban services, such as piped water from the local water authority, which makes living conditions expensive, unpleasant, and unhealthy.

Lack of servicing in the informal sector is not due to lack of demand for services on the part of residents. For example, water is essential and obtaining water privately is far more costly than the social user cost of public supply. While lack of servicing is sometimes characterized as a failure of governance, we hypothesize that such failure is often intentional. In other words, there is a strategic element to withholding service from informal housing sectors. Perpetuating bad living conditions for migrants is a way for existing residents to discourage in-migration to a locality. In China today, for example, this strategic component is explicitly articulated in policy design (Cai 2006). Forcing the vast majority of migrants into poorly serviced informal sector settlements is intended to restrain rural in-migration to China's urban areas, especially the largest ones. While explicit articulation of exclusionary motivations for withholding public services is less politically feasible in countries like Brazil, the data indicate that such policies were nevertheless adopted.

We study the determinants of servicing and its effect on migration to Brazilian urban jurisdictions from 1980 to 2000. While Brazil's rural to urban migration today has diminished, local resistance to in-migration was an important issue in the 1980s. In Brazil, localities within urban areas make policy decisions that influence in-migration to their specific localities, potentially providing an element of a race to the bottom within urban areas in terms of servicing migrants. We examine (a) the impact of withholding public water services to migrants on a locality's subsequent population growth and social composition, and (b) what determines a locality's provision of services to migrants. By studying Brazil's historical experience with urbanization and informal housing, we may better understand the forces at work in today's rapidly urbanizing countries.

An additional motivation for the study is that the development of unserviced informal housing sectors has effects beyond restraining migration, such as inequality in living conditions and unhealthy neighborhoods with high negative externalities. The resulting negative externalities may also affect the location decisions of those who live in nearby formal sector housing. An issue we do not directly study but that is relevant is that underservicing today requires later, and possibly more costly, catch-up investments. Installing water and sewer infrastructure long after the development of dense neighborhoods can require extensive spatial reconstruction and reconfiguration of neighborhoods.

There is an extensive literature on exclusionary policies of local jurisdictions in developed countries (e.g., the Tiebout literature as reviewed in Epple and Nechyba 2004). The reasons for exclusion in developing countries have the related elements of limiting over-population and congestion, as well as stratification, with the rich wanting to separate from the poor; but there are key differences between exclusion in developed and in developing countries. In developed countries, exclusion occurs through formal sector housing restrictions with the imposition of development fees, zoning, and other policy levers limiting housing development and population density. In many developing countries, however, formal sector housing restrictions, rather than halting locality growth, lead to the development of an informal sector where such restrictions are ignored. Unlike in most developed countries, the existence of an informal housing sector in developing countries is often tolerated. It is either not politically feasible to halt the development of informal settlements or weak institutions make it difficult to enforce laws against informal settlements. This paper takes the existence of weak institutions and enforcement as a given, at least historically, in Brazil. Because formal sector restrictions are not enough to halt growth, policies to exclude rely on withholding service to informal housing sectors.

Brazil has two types of informal housing markets. The first are *favelas*, which were historically created by invasions of government land or private land often under title dispute. In principle, such settlements are illegal, both because land-use regulations are evaded and because the housing is on land owned by parties other than the occupier. The second are *loteamentos*, where developments do not meet zoning regulations, but are built on legally acquired land. After development, however, owners cannot obtain land title because their housing does not meet zoning regulations. *Favelas* are an early phenomenon, often pictured in cities such as Rio de Janeiro as a response to in-migration pressure and lack of formal sector housing. Development of *loteamentos* was supposedly spurred by a national law in 1979 requiring 125 square meters of land as the minimum lot size for any construction (Avila 2006). Since only

15% of urban housing units in Brazil are apartments, the law was aimed at single family homes. A common view is that the law made formal sector housing unaffordable for many low- and low-middle income families at that time (Avila 2006 and Dowall 2006).

Until the late 1980s and democratization, it was in principle illegal for localities to provide public infrastructure such as central water connections in *favelas* and *loteamentos*, although some localities still did so. In other localities, the fact that *favelas* and *loteamentos* were illegal settlements provided an opportunistic excuse to deny or limit service provision. While localities cannot stop the construction of informal housing, they can contain the demand for it by withholding basic public infrastructure from informal neighborhoods, so as to offer poor living conditions and the need to substitute expensive private alternatives for services that could be provided relatively cheaply by the public sector.

Underservicing of informal neighborhoods may also arise because of discordant local and national government incentives. In developing countries, certain localities may be favored by policy initiatives of national governments, such as capital market allocations for industry, licensing for export, foreign direct investment, imports, and government investment in state capitalism. This political and economic agenda of the national government to favor certain localities has the side-effect of attracting migrants seeking job opportunities to these localities. If in-migration is unfettered, such favored localities ultimately become over-populated, with migration only slowing when the increased congestion, living costs, and diminished quality of life from over-population lead to dissipation of the benefits of national government favoritism. The literature makes this point generally (Ades and Glaeser 1995, Davis and Henderson 2003) and then with examples from Indonesia and China (Henderson and Kuncoro 1996, Jefferson and Singh 1999, Au and Henderson 2006), as well as Brazil.

In most developed countries, natural amenities rather than national government favoritism are the driving force of increased migration pressure to certain localities. In a recent paper, Gyourko et al. (2006) examine differences in natural amenities across American cities, making some key points. Certain superstar cities are favored with excellent natural amenities demanded more intensively by higher income individuals. In such cities, the authors observe that population growth over time has slowed, the share of population from higher income groups has grown, and there is excess demand to enter the city as evidenced by rapidly rising housing prices relative to the rest of the nation. The presumption is that superstar cities impose strict land-use regulations that inhibit further residential development which, if it occurred, would dissipate the natural advantages these cities offer. Higher income individuals are willing to pay higher housing prices associated with restrictions and better amenities, so that with national population growth, superstar cities become relatively richer. While national government favoritism may be more of an issue compared to natural amenity differentials in many developing countries, we also observe that high income localities are both slower growing and increasingly richer in Brazil.

The primary focus of our discussion and later modeling concerns localities' attempts to restrain in-migration as a way of avoiding over-population and dissipation of amenities. However, the empirics acknowledge that localities' attempts to limit in-migration may contribute to stratification of rich and poor households. Exclusionary policies may disproportionately target low-skilled migrants, who generate greater negative externalities and fiscal costs, with localities generally more willing to accept high-skilled migrants. Second, as in the superstar cities case, the localities most likely to implement exclusionary policies may be those that are richer and larger, and therefore have greater fiscal reasons (not wanting to subsidize poorer residents) or personal aversion to allowing low-skilled migrants to enter.

Section 2 of the paper discusses data and trends concerning stratification and exclusion in Brazil. Section 3 develops a conceptual framework to inform econometric specifications. Section 4 estimates the impact of servicing migrant households on locality population growth and composition. Section 5 analyzes whether localities in an urban area seem to interact strategically to exclude migrants and what types of localities are more likely to under-provide public infrastructure to migrant households. Section 6 concludes.

2 Urbanization and Public Infrastructure in Brazil

This section provides background information on Brazil relevant to our study. First we describe the data and spatial units of analysis. Then we provide an overview of Brazilian locality and urban area growth, which will help frame the precise approach and modeling we undertake. Finally, we examine data on different dimensions of housing sector informality and then turn to the issue of how, in the data, we represent policy initiatives that are based on exclusionary considerations.

The paper focuses on the post-1980 time period. We first examine the population response to public water service across localities, specifically how the 1991 level of public water provision to the houses in which low-skilled migrants are likely to live affects locality population growth and composition between 1991 and 2000. This timing turns out to be convenient in terms of an identification strategy. The 1980s are the last phase of Brazil's period of rapid industrialization and urbanization. Industrial development, which had focused on Sao Paulo and Rio de Janeiro in the post World War II period, starts to decentralize in the late 1970s with substantial and on-going industrialization of hinterland cities. This decentralization is facilitated by inter-city investments in transportation and telecommunications, as well as agricultural developments in the Northeast of Brazil (Da Mata et al. 2005). By the 1990s, these adjustments are largely complete. The change in urbanization and industrialization patterns pre and post 1990 reflect changes in the underlying drivers of city growth from the 1970s to the 1990s. We use this change in the determinants of city growth as part of the identification strategy in Section 4, in order to isolate the effect of public water provision on a locality's population growth and composition.

After establishing that public water provision affects in-migration, we turn to the determinants of public water provision. We look at provision during the 1980s, which starts off as non-democratic. Full democratization at the national level occurs in 1988, with democratic reforms in subsequent years.[2] One set of reforms removes restrictions on localities' provision of infrastructure services to the informal sector; another encourages the regularization of informal housing sectors and the upgrading of services. Our working assumption is that during the 1980s, exclusionary under-servicing of informal neighborhoods by localities was possible, even though most of our localities had elected mayors. Elitist dominated cities could legitimately deny services to the informal sector. However, by the 1990s and following national reforms, such exclusionary behavior was less politically feasible, something we observe in the data.

2.1 Data

We have Brazilian Population Census data for 1970, 1980, 1991, and 2000. These data 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. While we do not focus on land-use regulations, we do have retrospective information on regulations for each locality. A census of local governments conducted in 1999 and in 2005 indicates whether cities had passed a minimum lot-size zoning law in excess of the national standard of 125 square meters by 1999.[3]

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, we combine municipalities into common denominator ones, which we call localities or, more formally, Minimum Comparable Areas (MCAs).[4] These localities are constant spatial units during the time period we analyze. We focus on localities that are at least 50% urbanized by 1991 that are located in larger urban areas with multiple localities.[5] We drop from the sample all localities that form their own isolated urban area since we are interested in interactions among localities and since we wish to control for urban area fixed effects. This leaves a sample of 327 localities in 54 urban areas. Of these 327 localities, 238 are composed of just one municipality and 89 now have multiple municipalities. Of these 89, 45 have a dominant municipality with over 85% of the locality population. We will perform robustness checks where we drop the 44 localities with no dominant municipality or all 89 that have multiple municipalities. Additionally, since the urbanization rate of the 327 localities changes dramatically from 1970 to 2000, we often look at sub-samples of localities, imposing the restriction that they be at least 50% urbanized in a given census year in order to be considered as part of the urban area.

---

[2] The new constitution was signed in 1988. The government moved from military to civilian control in 1985, although without democratic elections.
[3] Data are from IBGE, Perfil dos Municipios Brasileiros, 1999 and 2005.
[4] Da Mata et al. (2005) provide a discussion of how to recombine Brazilian municipalities into Minimum Comparable Areas.
[5] These larger urban areas are roughly equivalent to Metropolitan Statistical Areas in the United States.

2.2 Patterns: Urban Growth and Stratification

In the spatial development of Brazil, urban areas experience mostly parallel growth from 1980-2000, with all but the two largest (Rio de Janeiro and Sao Paulo) growing at about the same rate, as a number of theories predict (Black and Henderson 1999, Gabaix 1999) and Figure 1a demonstrates. While the dispersion of growth rates is larger for smaller urban areas in Figure 1a, excluding Rio de Janeiro and Sao Paulo, there is no convergence or relative mean reversion. Urban areas grow mostly in parallel, with knowledge accumulation and improved education levels being the main drivers of urban area growth (Da Mata et al. 2007). In contrast, in Figure 1b, localities within urban areas experience much stronger mean reversion: bigger localities grow at a slower rate than smaller ones. From Figure 1b, it is also evident that the larger, slower growing localities tend to be richer.[6]

With population growth in an urban area, old localities fill up and become crowded, and new localities develop. With economic development, urban areas spread out, fueled by declining commuting costs and transportation improvements that make central city locations less valuable. This movement of migrants into different localities within urban areas is the variation we will utilize in the empirical work. In Brazil, as in much of the world, the rich live predominately in the center cities and the poor live in suburbs. Much of the exclusion we observe is therefore of over-crowded central cities and richer and larger suburbs deflecting migrants into low-income, suburban localities within the urban area.[7]

Finally, localities in Brazil also appear to subscribe to Gyourko et al.'s (2006) superstar city story, with rich localities becoming even richer over time. In Figure 2, we plot the share of a locality's households that is among Brazil's richest 10% in 2000 against the share that is among the richest 10% in 1980. Here, we note that the largest 20% of all localities have almost universally increased their share of the rich, with their data points lying above the 45-degree line.

2.3 The Informal Sector and Public Infrastructure

An issue is how to identify who lives in the informal sector and how to define that sector. We summarize some possibilities as discussed in the literature (Dowall 2006, Biderman 2007). In the census, there is a question filled out by census takers 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.

---

[6] Rich localities are defined as those that are in the top 20% of all urban localities sorted by median household income.

[7] Why the difference in where the rich live compared to the United States? One reason may be that, unlike in the United States, in most countries, funding for public education occurs at the state or national level. The rich by suburbanizing cannot form exclusionary clubs offering independently funded, high-quality schooling; and thus they may prefer the center city with its lower commuting times to service-intensive central business districts.

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 do not report land title. Again 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 eviction even from *favelas* is rare. As an example of the perception of security, in regularization programs to grant land title to those without it, some participants fail to take the last step (about one day of work) and register their land tenure once they are able to do so.

A different approach is to define informality based on lack of public infrastructure provision. The literature (see Dowall 2006) 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. There are several reasons we focus on water connections instead of full service. First, the literature argues that it is lack of water which is the key issue for residents (e.g., see Scheper-Hughes 1993, Chapter 2). Second, provision of a central water connection appears to be more of a locality decision made by municipal water authorities or in negotiation with regional authorities, while sewer provision seems to be more of a state-level decision. Third, many neighborhoods, even richer ones historically without sewer connections, continue to rely on private alternatives such as septic systems. In 1991, only 187 localities from our sample of 327 localities had more than 10% of houses with full service. For water, 326 of 327 localities provided a central water connection to at least 10% of houses. If no households in a locality have sewers, it is not an exclusionary tool.

For water, lack of a public connection means private alternatives must be used. In many localities, especially those situated on large water tables, the private alternative for richer households in the early years of urbanization was to dig wells, and some cities initially chose to have no central water provision. However, with sustained use and population growth, wells in many areas started to run dry for portions of the year, leading to the development of central water systems in almost all localities. Low income migrants cannot afford to drill a well, even if these remain a viable alternative to public water provision. For migrants without a central water connection, the private alternatives at the margin are to use a public stand pipe and haul the water for some distance, extend hoses into nearby but often polluted rivers and pump the water to their homes, subscribe to water truck services (known as *carros-pipa*), or purchase bottled or bagged water. These are difficult, dirty, or expensive alternatives to central 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 cities of Sao Paulo and Rio de Janeiro. The next section describes how these hedonics fit into the overall conceptual and estimation framework. 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 in Appendix A. The regressions are for 1980, the one census year in which relevant data on rents and neighborhood location within the two cities are available, so we can use neighborhood fixed effects to control for neighborhood characteristics. We find the marginal consumer is willing to pay 12% more for a rental unit with a central water connection (net of any additional premium for indoor plumbing) in Sao Paulo and 23% in Rio de Janeiro. The regressions control for types of sewage provision and whether households have electricity.[8]

In this paper, we focus on the notion of exclusion through lack of servicing, but how does a policy of restricted servicing appear in the data? This is a complicated issue because in Brazil (unlike China) exclusionary policies cannot target individuals based on personal characteristics such as migrant status. They can only target houses and neighborhoods where migrants are likely to live, recognizing that migrants can live where they demand given incomes and prices. We start by looking at the evolution of servicing as a guide to how we will measure the extent of servicing that migrants are likely to face.

2.4 Provision of Infrastructure Services

Table 1 explores dimensions of water servicing. We look at localities that are at least 50% urbanized by 1991, as well as localities that are at least 50% urbanized versus those that are not in 1970, allowing us to track service expansion within constant samples of localities. Table 1 shows the rapid expansion in services in urban Brazil over the decades. In 1970, localities that were at least 50% urbanized provide a connection to a central water system to 50% of their urban households. By 2000, this number has reached 89%, reflecting growth along two margins – increased servicing within highly urbanized localities and the addition of localities that were previously less than 50% urbanized. The fact that the weighted percentage of houses with a central water connection is higher than the unweighted percentage suggests that more populous localities service a greater share of their households.[9]

The second panel 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 localities dramatically increase the share of houses they service, from a mean of 50% in 1970 to 92% by 2000. Finally, in the third panel, we explore the increase in

[8] We also note that renters are willing to pay an additional 20% in Sao Paulo and 36% in Rio de Janeiro for a unit with central sewer and electricity, in addition to water. Isolating the components in Appendix A, there is a very high premium on electricity (although even in 1980 it is virtually universally available), but central sewer itself still commands a 9% and 18% premium in Sao Paulo and Rio de Janeiro, respectively, over no connection (with septic systems in those congested cities generating little premium over no sewerage at all).

servicing among today's urban localities that were under 50% urbanized in 1970. These newer localities also have rapid expansion, but are at lower service levels. The numbers suggest that localities can quite rapidly expand their central water systems, and the fact that some houses remain unserviced even by 1991 and 2000 might reflect deterrence of in-migration.

The first panel of Table 1 also explores how services differ across types of households in 1991. In terms of tenure, not surprisingly, those who report not owning the land under their house are less serviced than those who report land title. Small houses, those with 1-3 total rooms, are less serviced than large houses, with 7-9 total rooms. While migrants are less serviced than non-migrants, we try to control for income effects by looking at households in the bottom 20% of the national urban income distribution. Low-income migrants (who moved to the locality in the last ten years) are less serviced than low-income non-migrants, although the differences within low-income groups are not large. While localities care about the rate of in-migration, today's migrants are tomorrow's non-migrants. Localities may be concerned about the level and extent of poverty in their population, potentially welcoming higher-income migrants (who may displace existing lower-income migrants) and discouraging lower-income migrants from entering. In Section 4, we will look at both growth and population composition outcomes.

While we can cite numbers for servicing of migrants, as noted above, localities cannot discriminate on the basis of income nor migration status. What localities can do is not service the houses that most migrants and low-educated households are likely to occupy. We look at this in Table 2. Before turning to the table, we note that migrants tend to live in the smallest houses, defined as those with 1-2 total rooms in 1970 and 1980 or 1-3 total rooms in 1991 and 2000. In contrast, we also look at households living in larger houses, those with 6-7 rooms in 1970 and 1980 and 7-9 rooms in 1991 and 2000 (groupings which exclude approximately the top 10% of houses by size in each decade). In the data, overall, about 15% of households live in small houses and 20% live in these larger houses, with exact numbers footnoted in the table. In 1991, 33% of the migrants in our localities lived in small houses and only 12% lived in large houses, whereas 18% of non-migrants lived in small houses and another 18% lived in large houses. For migrants from rural areas, 42% lived in small houses and only 7% lived in large houses. Finally, for low-educated households (where the household head did not complete primary school), 27% lived in small houses and 13% lived in large houses, compared with 17% in small houses and 19% in large houses for higher-educated households (where the household head completed at least primary school). Targeting the smallest houses therefore appears to be a mechanism to discriminate against migrants and low-educated households.

In Table 2, we examine the service levels for small versus large houses. Table 2 gives locality averages for servicing of small versus large houses for the full sample of 327 localities, for the restricted

---

[9] We weight the percentage of houses serviced with a central water connection by the number of urban households

sample of 185 localities that were at least 50% urbanized in 1970, and for the remaining 142 localities that were less than 50% urbanized in 1970. For the full sample of 327 localities, only 48% of small houses in 1980 had a public water connection, compared to 74% of large houses. By 2000, the gap diminishes but is still noticeable overall. Restricting the sample to only those localities that were more or less than 50% urbanized in 1970, we observe that the gap in servicing of small versus large houses is similar to what we found when looking at the full sample, but the absolute service levels of both small and large houses are higher for those localities that urbanized earlier and lower for those localities that urbanized later. In Sections 3 and 4, we will use the provision of public water connections to small houses as our basic exclusionary measure, representing the quality of infrastructure that incoming migrants might expect when choosing to live in a particular locality.

2.5 Land-Use Regulations

As an interesting aside, we examine one aspect of local land-use regulations: lot-size zoning over and above the national 1979 minimum lot-size law. Most of these local regulations were passed after democratization in 1988, and our data are from 1999 and 2005. From our sample of 327 localities, 200 had passed a minimum lot-size zoning law in excess of 125 square meters by 1999 – that is, a minimum lot-size zoning law in excess of the national standard. Comparing these localities, when we look at homeowners without title to land relative to those with title to land in zoned versus unzoned localities, we obtain a ratio of 1.06. In other words, the ratio of homeowners without title relative to those with title is slightly higher in zoned versus unzoned localities. On the other hand, zoned localities have slightly less in-migration, so that the share of migrants to residents in zoned versus unzoned localities is 0.97. These numbers suggest tougher zoning is associated with both reduced overall growth and increased informal sector development. Presumably, the adoption of tougher zoning is endogenous. Those with tougher zoning may have faced potentially higher growth, and therefore imposed tougher regulations to curtail it. The general issue of endogeneity will be critical in the later identification of the effects of servicing on locality growth.

3 Conceptualizing Exclusionary Behavior

This section develops a model upon which we base the key aspects of the empirical formulations examining locality population growth, service levels, and strategic interactions across localities. The empirical work is focused on variations across localities within urban areas for identification of effects. For much of the work, this means we can ignore the determination of urban area characteristics, which are captured by urban area fixed effects. For example, we assume workers in all localities in an urban area

in the locality.

participate in the same overall urban area labor market. Then, conditional on total urban area size, people's choice of locality within an urban 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 are choosing 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 urban area faces a supply of in-migrants, which will be split across the localities of the urban area depending on the living conditions in these localities. Incumbent residents of localities value better services for these migrants for either or both altruistic or externality reasons, which is a force to increase service levels. However, they prefer fewer migrants to their own locality, which is a force to reduce service levels. For economic growth reasons, incumbent residents of a locality may want more migrants overall to the urban area as long as these migrants do not live in their own locality. 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 urban 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, we allow each locality to choose a level of servicing. While we model the equilibrium in a one shot context of a city facing a potential wave 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 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 neighborhoods, where they must privately secure services at a higher unit cost, $c > c_0$ (e.g., water purchased from water delivery trucks). By living in the informal sector we assume migrants escape locally set property taxes, paid by incumbents.

3.1 *Equilibrium within the Informal Sector*

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

$$\left.\begin{aligned} U_0 &= U_0(w(\bar{N}+\bar{L}), p_0, c_0, b, L, Z) && \text{(a)} \\ h_0 &= h_0(w(\bar{N}+\bar{L}), p_0, c_0, b, L, Z). && \text{(b)} \end{aligned}\right\} \quad (1)$$

Disposable income of migrants is $w(\bar{N}+\bar{L})$ which could vary with overall urban area scale, as measured by total incumbents in the urban area, $\bar{N}$, plus total migrants to the urban area, $\bar{L}$. In the locality, p_0 is the price of housing in serviced neighborhoods and c_0 is the unit cost of water, while b is the share of migrants who are serviced in the locality, potentially a positive externality for both migrants and incumbent residents. For example, less servicing means more people using and disposing of untreated water which could result in health externalities. The share of migrants served is a policy variable influenced by the locality. The endogenous number of migrants to the locality itself is L, which may also generate congestion or other externalities. 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 demand functions for housing and an indirect utility of the form

$$
\begin{aligned}
U &= U(w(\bar{N}+\bar{L}), p, c, b, L, Z) && \text{(a)} \\
h &= h(w(\bar{N}+\bar{L}), p, c, b, L, Z). && \text{(b)}
\end{aligned}
\qquad (2)
$$

For the same housing price p, $U_0 > U$, given $c > c_0$. To equalize utility across the two types of neighborhoods within a locality requires $p_0 > p$, and from indirect utility in equations (1a) and (2a), we have

$$p_0 = p_0(p, b, c, c_0, L, Z; \bar{N}+\bar{L}). \qquad (3)$$

Equation (3) underlies the hedonic regressions reported in Appendix A for Sao Paulo and Rio de Janeiro, examining within locality differences in relative rents based on type of service. 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 urban area. We turn to this next.

3.1.1 Housing Demand Equals Supply

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 number of serviced sites in the informal sector is a policy variable, so the supply margin 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 equation (1) and of the $L - L_0$ unserviced people in (2), and using equation (3) for p_0, we have

$$L\left(b\cdot h(w(\cdot), p_0(\cdot), c_0, b, L, Z) + (1-b)\cdot h(w(\cdot), p, c, b, L, Z)\right) = H^s(A, p), \quad \text{(a)}$$
$$b \equiv L_0 / L. \quad \text{(b)} \qquad (4)$$

In equation (4b), $b \equiv L_0 / L$ is the proportion of migrants served, the basic policy variable.

3.1.2 Supply of Migrants

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

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

Substituting for p from equation (5) into (4a), we get

$$L = L(A, c_0, c, b, Z; \overline{N}+\overline{L}). \qquad (6)$$

Using urban area fixed effects (conditioning on $\overline{N}+\overline{L}$), we will estimate a version of equation (6) to show how the policy instrument, b, as well as housing supply conditions, A, affect locality population. Although service-level differences within localities are capitalized into intra-community differences in housing prices, locality choice by migrants is affected by relative service levels. Servicing affects overall housing demands and hence price levels across localities.

3.2 Equilibrium across Localities within an Urban Area

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

$$L_{ij} = L_{ij}(\mathbf{b}_j, \mathbf{A}_j, \overline{N}_j, \overline{L}_j, \boldsymbol{Z}_j, \mathbf{c}_{0j}, \mathbf{c}_j), \qquad i = 1, \ldots, n_j, \qquad (7)$$

where n_j is the number of localities in urban area j, $\mathbf{b}_j$ is the vector of service ratios of localities in urban area j; $\mathbf{A}_j$ is the vector of land supply endowments for migrants in each locality in urban area j; and $\boldsymbol{Z}_j$ is the vector of other characteristics. Unit costs of services could vary by localities within an urban area, in which case we would have vectors, $\mathbf{c}_{0j}$ and $\mathbf{c}_j$. Finally, the function is also indexed since its form will vary with the number of localities, n_j.

Given $L_{ij}(\cdot)$ functions, we can calculate the population response of any locality in an urban area to a change in another locality's policy variable, b_{ij}. Thus, if we are looking at locality 1, we can calculate $\partial L_{ij} / \partial b_{1j}$ for all $i \in j$, which is required below in assessing potential strategic responses.

3.3 The Choice of Servicing for Migrants

We start with the notion that incumbent local residents of locality 1 of a representative urban area choose b_1 to maximize utility. There are two issues here that we need to address. The first is under what conditions incumbent residents are free to choose b_1. By incumbents we mean there are a set of wealthier long terms residents in each Brazilian locality in the 1980s who have sufficient control over the political process and selection of mayors so as to strongly influence exclusionary policies, including servicing of housing in migrant neighborhoods. In principle, even if local water supply authorities operate at a more regional level, localities can determine or strongly influence where water pipes are laid. However, with democratization, localities came under strong pressure to make piped water universally available. The second issue is the extent to which localities interact strategically. We explore both these issues in the empirics. For now we proceed as though localities have complete control over water supply and interact strategically.

We use a reduced-form specification of preferences,

$$V(y(\bar{N}+\bar{L}), \tau_1(b_1 L_1;\cdot), \tilde{c}_1, b_1, L_1, Z_1), \qquad V_y, V_{b_1}, V_{\tau_1}, y_{\bar{L}} \geq 0,\ V_{L_1} < 0. \tag{8}$$

For the first argument, we allow more migrants, $\bar{L}$, to the overall urban area to increase incomes of urban area residents, $y(\bar{N}+\bar{L})$. This models net scale effects in urban area total employment. The fourth term, where $V_{b_1} > 0$, reflects positive externalities from better servicing of migrants in the locality. However, the fifth term, where $V_{L_1} < 0$, implies that while incumbent residents of locality 1 want more migrants to enter the urban area, they do not want them in their own locality. This could reflect local congestion considerations or prejudices against migrants. Z_1 are other characteristics of the locality.

The second and third terms in (8) are important. We assume incumbents are owner-occupiers of fully paid off housing and (8) includes the continuation value function for future periods (where the future is expected to be the same as today). Alternatively, this is the final period of the model. $\tau_1(\cdot)$ is the current housing cost of incumbents. This housing cost incorporates incumbents' 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 units, $b_1 L_1$, and is also a function of past servicing. For example, if more houses were connected historically, new houses may be cheaper to service (e.g., just extend existing mains versus install entirely new systems). Incumbents may have public water connections so that $\tilde{c} = c_0$ or they may have wells in which case $\tilde{c}$ may be close to zero.

We assume for the moment that incumbents in locality 1 interact strategically with other localities. Incumbents choose b_1 to maximize equation (8), holding other localities' choices of service levels fixed (the Nash assumption), accounting for how the choice of b_1 affects other localities' migrants. Maximization gives

$$V_y y_{\bar{L}} (\sum_{i, i \in j} \partial L_{ij}(\cdot) / \partial b_1) + V_{\tau_1} (L_1 + b_1 \partial L_{1j} / \partial b_1) + V_{b_1}(\cdot) + V_{L_1}(\cdot)\ \partial L_{1j} / \partial b_1 = 0, \qquad (9)$$

where the $\partial L_{ij} / \partial b_1$'s are calculated from equation (7). Using this first order condition and an equation (7) for each locality in the urban area, we can define

$$b_{1j} = b_{1j}(\mathbf{b}_{-1j}, \mathbf{A}_j, \mathbf{Z}_j, \mathbf{c}_{0j}, \mathbf{c}_j; \bar{N}_j, \bar{L}_j), \qquad (10)$$

where $\mathbf{b}_{-1j}$ is the vector of service levels in other localities in the urban area. Equation (10) will be the basis for estimating how characteristics of existing localities influence policy choices and for testing strategic interactions. For strategic interactions, by differentiating equation (10), we can solve for db_{1i} / db_{ij}, which yields how locality 1 changes its service levels in response to a different service offering in locality i. Perhaps of greater interest empirically, as suggested in Figures 1b and 2, will be how different localities set service levels for migrants according to whether they are higher-income versus lower-income or larger versus smaller.

3.4 Extensions

So far we have looked at migrants assuming they are generally low-skilled, and we have assumed existing residents are high-skilled and immobile. In actuality, we may also have in-migration of high-skilled individuals from outside the urban area as well as movements of existing high- and low-skilled residents across localities. While we will estimate overall locality household growth equations based on equation (6), we will also separately estimate growth equations for high-skilled and low-skilled households. We introduce other considerations, such as differing tastes across localities among existing residents concerned with inequitable provision of public services, as reflected by their voting preferences.

4 Effect of Service Provision on Locality Growth and Population Composition

What is the effect of servicing decisions on locality population growth? The specification is based on a linearized version of equation (6) with urban area fixed effects. Controls such as total migrants to the urban area, local servicing standards, and urban area wages are swept into the fixed effect. What we examine is the within urban area allocation of migrants across localities. The basic estimating equation is

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

We look at locality population growth between 1991 and 2000 as a function of locality characteristics in 1991. From equation (7) these include the level 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.

A key issue in estimation concerns the error structure. The urban growth literature (e.g., Glaeser et al. 1995) often takes the stance that (a) covariates are pre-determined and not affected by contemporaneous shocks that might induce growth, and (b) 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 predetermined 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 current housing supply conditions today.

We focus on locality growth from 1991 to 2000. This last interval in the census data allows us to separate Brazil's initial rapid industrialization and urbanization that occurs after World War II and extends into the 1980s, from today's more modern economy. By 1991, Brazil is 75% urbanized. The axis of industrialization that had focused on Sao Paulo and Rio de Janeiro in the Southeast of Brazil expanded with substantial and on-going industrialization of hinterland cities and more rapid growth in the Northeast region. The drivers of local growth have changed with the development of new export markets and the development of new agricultural crops for export, as well as the move from heavy industry based on state capitalism to lighter industry based on the manufacture of consumer products (Da Mata et al. 2005). This change in economic regimes will be part of an identification strategy. We note that despite the high level of 1991 urbanization, there remains on-going migration as well as population growth. In our sample, the number of urban households grew by 40% from 1991 to 2000. We look for growth effects in the democratic era to see if poor servicing in 1991, which arose in the non-democratic 1980s as explored in the next section, affects growth from 1991 to 2000.

4.1 Instruments

We need to instrument for two types of variables. First is the service variable: public water provision to migrant households. Second are housing supply conditions in the locality. Given that we have urban area fixed effects, instrumenting is challenging since we need to predict intra-urban area variation in covariates. While we experiment with a parsimonious instrument list, we ultimately rely on a longer list to get more power.

4.1.1 Extent of Servicing

For water supply, we turn to the inferred history of water provision in the locality, which influences the costs of extending central water provision today. Ideally, we would like 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. There is also the issue that with wells incumbents have an almost zero marginal cost of water and have limited demand for central provision for themselves. We unfortunately do not observe wells in the data as a separate item, however we know the geological 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 migrants and more generally 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 (trucks, bags, bottles). Thus, we believe that instrumenting for servicing using geological and weather variables meets the exclusion restriction: these instrumental variables are correlated with public servicing levels, but have no direct, independent effect on current household growth rates. A caveat is that equation (6) contains c and c_0. These could be influenced by depletion of overall water resources from use of wells. Our assumptions are that urban area fixed effects help condition on area wide depletion and that what is relevant is the ratio c / c_0 (e.g., the cost of hand-delivered water relative to the cost of publicly provided metered water), a ratio that may not vary much within an urban area.

What geological conditions determined the efficacy of wells historically? 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 (IUSS Working Group WRB 2006). Second, insolation affects the rate and variability of underground water replenishment, again a key issue in the viability of wells (Markstrom et al. 2008). Insolation is a measure reflecting the amount and intensity of sunlight reaching the earth's surface. High insolation is associated with less rainfall and 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. While insolation does vary across localities in an urban 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 porous geology (reducing the need for servicing, historically), porous geology interacted with mean insolation (increasing the need for servicing), and 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 are strong instruments for servicing, with the expected effects.

4.1.2 Housing Supply Conditions

The other 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 are potentially the number of urban households, average education (influencing the demand for space, as well as being a potential amenity), and the share of households that are rural in the locality (influencing the potential supply of higher density urban housing). The main instrumental variables strategy is based on two notions. First, unobservables that affected urban area and locality growth in the past are different from unobservables affecting growth today, so error drawings from 1970 are uncorrelated with drawings in the 1990s. As discussed previously, within urban areas, locality economic bases have changed (an unobservable not captured in equation (7)), as has the urban area labor market demand for the skill sets of households living in different localities. In addition, with economic development, there has been decentralization of economic activity within urban areas as well as to hinterland cities. The second notion is that past drawings affected housing and other irreversible investment decisions as well as locality population educational composition in the past. Historical accumulations are relevant since any adjustments away from them in locality characteristics are slow, so historical variables are strong instruments for 1991 covariates. If, in the 1960s, a locality attracted low-educated 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 (Z's) influences 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.

Use of historical instruments presents some complications. There is a tension between going further back in time to break the persistence in relevant unobservables and weakening the strength of the instruments. Apart from specification tests, it is difficult to prove that the assumptions are correct — i.e., this is at best a very limited version of a natural experiment. 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 time period, one reason why we focus on the latter time period for population growth equations. For instruments, we include the following variables either on their own or interacted with other instruments: (1) access of a locality to Sao Paulo markets, which played a critical role historically, before the development of modern trans-national transportation systems even though today it has little impact on growth;[10] (2) the illiteracy rate among the adult population in the locality and the rest of the urban area in 1970, which influences, through accumulation, the average educational attainment today; (3) the manufacturing-to-service employment ratio in the rest of the urban area in 1970, which helped urban area economic attainment at the time and influences local economic composition today; (5) the number of households in the rest of the urban area, which gives a historical size measure influencing urban size today; and (6) the share of the rest of the locality households that were rural 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 urban 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 socio-economic status of the locality population, in addition to geological conditions, will have influenced the use of wells in one locality compared to another in the past.

4.2 Effects of Servicing on Growth of Urban Households

Table 3 contains the basic results of the effects of servicing on the growth in the number of urban households. Columns 1 and 2 contain OLS estimates, with shorter and longer lists of covariates. Columns 3 and 4 report corresponding Two-Stage Least Squares (2SLS) estimates. 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.

4.2.1 Servicing

In OLS estimation in columns 1 and 2, the coefficient on servicing is strongly 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 strongly negative coefficient and reverses its sign, making it positive. This positive coefficient is large. For a point estimate of 0.74 in either column 3 or 4, a one standard deviation (0.21) increase in servicing leads to an increase of 0.16, or approximately 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-educated 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 urban area level). With shorter instrument lists the coefficient is more noisily estimated. For example, with just the porous geology variable, the geology variable interacted with the two insolation measures and urban population in 1970 in the rest of the urban area as instruments, in column 3, the coefficient for servicing is almost the same, 0.72, but the standard error is now 0.50. For the longer instrument list, the partial F-statistic on the first-stage regression for service levels is around 13 and the p-value for the Anderson Likelihood Ratio test statistic is 0.11, suggesting we cannot reject weak identification. We are instrumenting for multiple interrelated variables under urban 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 with 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.[11]

We perform two sets of robustness checks for the servicing effect in the growth equation. First,

---

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

[11] Estimating column 3 of Table 3 using LIML instead of 2SLS, we obtain estimates for the coefficient (standard error) on the servicing variable of 1.261 (0.648). When estimating column 3 of Table 3 using LIML and a shortened instrument list, we obtain estimates of 0.858 (0.506). 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 urban area in 1970, this ratio interacted with the log of the distance of the urban area to Sao Paulo, the log of the number of urban households in the rest of the urban 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. When estimating with LIML and the shortened instrument list, the p-value for the Anderson Likelihood Ratio test statistic is 0.03, suggesting we can reject weak identification.

we reestimate the model in columns 3 and 4 of Table 3, separately dropping (a) all 89 localities in which there are multiple municipalities, (b) the 44 localities in which there are multiple municipalities where the dominant municipality has less than 85% of the locality urban population, and (c) all localities in the two largest urban areas, Rio de Janeiro and Sao Paulo. The estimates do not change noticeably, so the inclusion of these localities has no impact on results. Second, as a service variable, we tried the share of houses without land title that have a central water connection, but our instruments are very weak. While the coefficient on servicing is consistent with the Table 3 result, it has a large standard error.[12]

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 -0.026 becomes -0.112 under 2SLS estimation; and for land (alleviating crowding), the OLS coefficient of 0.006 becomes 0.060 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.

4.3 Composition Effects

So far we have looked at how withholding service leads to a decline in overall in-migration. Because it is aimed at small houses, withholding service also acts to discourage in-migration of lower-income residents from other parts of the urban area as well as of new arrivals to the urban area. Thus, we should see withholding service negatively affecting the growth of lower-income households in a locality, which we represent by education level of the household head. We use education rather than income to better represent permanent socioeconomic status.

We define two groups: households where the household head has not completed primary school, which is about one-third of the urban population in our sample of localities in 2000, and those who have primary school or more. While we expect that poor servicing of small houses will affect the growth of low-education households in a locality, there is the possibility it will also detract from the growth of higher-education households. In the modeling in equation (8), we postulate that servicing of migrants is a positive externality for incumbent residents. Thus, we examine separately the effect of servicing small

[12] The best case involves the short instrument list used in the LIML specification (see previous footnote), where the first-stage partial F-statistic for the service variable is 6.20 and the p-value on the Anderson Likelihood Ratio test

houses on the growth in the number of lower-education and higher-education households.

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 lower-education household growth, using both OLS and 2SLS estimation, and columns 3 and 4 deal with higher-education household growth, again using OLS and 2SLS estimation. The deterrent effects of higher density and own group size are strengthened in moving from OLS to 2SLS, as expected, although the coefficients are not always significant.

For the effect of servicing small houses on the growth in the number of lower-education and higher-education households, as expected, and as in the overall growth estimation, OLS coefficients on the servicing of small houses are negative. With 2SLS estimation, the service coefficient for the growth of low-education households becomes positive, very large, and significant. 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.

As robustness checks, for higher-education households, we include a control for the servicing of large houses occupied by high-education households (not shown), which leaves other results unchanged and yields an insignificant coefficient on the large-house variable. When we estimate a ratio model (also not shown), where the ratio is the growth rate of low-education relative to high-education households,[13] the coefficient on servicing of small houses is negative and not statistically different from zero (the coefficient is -0.402, with standard error 0.338). Both the results from the ratio model and the results in columns 2 and 4 hint that poor servicing of small houses has adverse effects on the growth of all education households, suggesting a negative externality for higher-education households.

5 Determinants of Locality Infrastructure Servicing and Land-Use Regulation

In this section we examine how localities choose service levels, focusing on the implementation of equation (11). The empirical results will support the notion that localities in the 1980s under-serviced

---

statistic is 0.63. For this specification, the OLS coefficient of -0.164 becomes 1.821 under LIML but the standard error is 2.215.

[13] 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 γ 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: $[\ln(N_{i,t}^L) - \ln(N_{i,t-1}^L)] - [\ln(N_{i,t}^H) - \ln(N_{i,t-1}^H)] = \beta_0 \ln(N_{i,t-1}^L / N_{i,t-1}^H) + \beta_1 \ln density_{i,t-1} + \gamma \ln(b_{i,t-1}) + \varepsilon_{i,t}$.

migrants, potentially to deter in-migration. However the results do not support the idea that localities within a metropolitan area interact strategically, although our test of strategic interactions has limitations. In terms of limitations, we note that the local public finance literature and the industrial organization (IO) literature have diverged on how to examine strategic interactions. 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. In a local public finance context with the type of model we have outlined, there are two problems in following the IO approach. First, our model does not rely on a simple cost and demand function. There are many more moving parts or 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. Results therefore depend even more on functional assumptions. Second, our characterization of decision-making on service levels for migrants is itself a simplified reduced form specification, assuming incumbents set service levels. For these two reasons, we follow the local public finance literature and test for strategic interactions directly (as in Case et al. 1993 and Besley and Case 1995), in a reduced form context.

The local public finance approach has two problems. The first (which would be the IO critique) is that we can only identify strategic interactions locally, in the sense that the literature uses linearized specifications to estimate equation (10). Even the most simplified specification of primitives does not yield a linear form, and in typical functional specifications of tastes and technology, the relationships governing b_1 will be highly non-linear. Second, across urban areas, there are different numbers of localities; this could generate different forms to the b_1 equation, an issue ignored in the local public finance literature and one that does not exist in the IO literature with just one market.

A standard formulation to equation (10) based on the public finance literature as reviewed in Brueckner (2000) is

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

where $\sum_{i\neq 1,\ i\in j} w_{1i} b_i$ is a weighted sum of all other localities' choices in urban area j. Writing equation (12) for all localities, we have

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

where $\mathbf{W}$ is the weighting matrix, with zeros for the own locality and for all other localities not in the same urban area as the own locality. Weights for localities in the same urban area are normalized to sum to 1.

Weights are typically chosen based on no explicit model. However, since equation (10) 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.[14] 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 greater 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. We also report results using equal weights for all localities within the urban area.

Estimating equation (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, $\sum_{i\neq 1,\ i\in j} w_{1i}b_i$ is correlated with ε_1. Thus, if reaction functions are positively sloped, a shock which causes one locality to raise its servicing will lead others to raise their servicing. The absolute impact of reactions terms would be overstated. There are several possible solutions, including use of lagged b_i (as in Hayashi and Boadway 2001) and IV estimation (as in Fredriksson and Millimet 2002) where we could instrument for $\sum_{i\neq 1,\ i\in j} w_{1i}b_i$ with $\sum_{i\neq 1,\ i\in j} w_{1i}Z_i$, using time-lagged values of Z_i. Even though we instrumented for individual localities' b_i previously, instruments for $\sum_{i\neq 1,\ i\in j} w_{1i}b_i$ are too weak to support the IV approach. Therefore we use the third and most common approach. We rewrite equation (13) as

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

However, equation (14) is also flawed, because it ignores spatial correlation of the error terms, which almost certainly arises here, for example, from unobserved, correlated geographic factors across localities within an urban area that affect public infrastructure choices. Thus, following the literature, we assume an error structure of the form $\boldsymbol{\varepsilon} = \Psi\mathbf{M}\boldsymbol{\varepsilon} + \boldsymbol{\xi}$, where $\boldsymbol{\xi} \sim N(0, \sigma^2\mathbf{I})$ and $\mathbf{M}$ is a matrix of spatial weights. For this specification, the estimating model becomes

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

[14] To see how land-supply weights arise from the model, assume the following functional forms: preferences of migrants to locality i are $x + g^{\gamma} + b_i^{\delta}$, supply of migrants to the urban area is denoted by $\bar{L}$, land supply is $p_i = L_i A_i^{-1}$, 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 in locality 2.

For the $\mathbf{M}$ matrix, we use weights calculated from the inverse distance between pairs of localities in an urban area (normalized to sum to one), which gives greater weight to neighboring localities with similar geography. In estimation of (15), for $\mathbf{Z}$ and $\mathbf{M}$, we use lagged covariates to deal with issues of contemporaneous correlation between errors and covariates; however, policy choices, $\mathbf{b}$, are contemporaneous.

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. We estimate different specifications of the model, starting with a base case where we try to determine a reasonably robust econometric specification, and we explore counterfactuals. We estimate the model with and without strategic interactions. The model with strategic interactions is equation (15); the model without strategic interactions is an OLS version of equation (12) without the $\sum_{i\neq 1,\ i\in j} w_{1i}b_i$ term but generally with urban area fixed effects. In (15) we cannot have urban area fixed effects since by construction that biases the coefficient γ to be negative, which can be seen by inspection for the case where an urban area has just two localities. Since we find no solid evidence of strategic interactions, we begin by presenting results where they are not included. Note that strategic interactions state that incumbents in locality i account for the effect of other localities' choices of servicing b_k, $k \in j$ and $k \neq i$, on their b_i choice, and there may be several reasons for why we do not observe localities strategically interacting. Localities may not know the b_k of other localities with any degree of precision and may have limited ability to respond as those b_k change. A locality is not a firm instantaneously adjusting its price in response to price changes of other firm; changes require alteration of infrastructure plans and use of infrastructure funds. Given the slow, local political processes by which incumbents influence b_i, allowing for more sophisticated behavior may not fit the data.

The base case without strategic interactions is in Table 5. 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 urban 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. Second, the urban literature hypothesizes that there are scale economies in public service provision which would lead larger localities to provide services more cheaply, although such scale effects may disappear at modest sizes. However, we also anticipate that larger, richer localities may have more incentives to deflect migrants. So while they value *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-educated 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 (in terms of public service provision) effects, we expect the interaction between these two to be negative. The idea that these effects are specific to servicing of housing geared to migrants, and hence by inference are concerned with deterring migration will be reinforced by a look at two counterfactuals later.

A longer covariate list adds in one taste related and two cost related factors. The taste factor represents preferences for more egalitarian policies, taken as the share of voting in the locality in favor of anti-military political parties in the 1982 elections for representatives to the national legislature. This is intended to measure preferences for more leftist representatives who might be more egalitarian (as revealed by actions in the subsequent democratic era). One cost factor is higher density, where greater density would entail stronger negative externalities from poor water and sanitation conditions and also reduced cost per household of laying water mains. The other cost factor is the overall level of servicing in the locality in 1970, driven by geologic conditions and 1970 factors during a different political regime. Higher prior servicing represents enhanced scale effects—less cost to extend existing mains rather than installation from scratch.

Base case results are in Table 5. Column 1 gives the results with the short covariate list and 2 with the longer covariate list. We focus on the column 2 results. 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 0.11, while at one standard deviation above mean size (10.25), the same increase in income leads to an increase in servicing that is almost 50% smaller (at 0.05). 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) does not change servicing, while at one standard deviation above mean income (at 9.91), the same increase in size leads to a *decline* of servicing of 0.05 (from a mean of .77).

Table 6 explores this negative income-scale interaction further. We divide localities into separate size and income quintiles and then interact these, creating 24 cells relative to the base. Results differ somewhat for the column 1 and 2 specifications of Table 5 in terms of interactive effects, but the pattern is clear in both. As Table 6 shows for the column 2 specification, 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. Interactions are always negative, but, with one exception, are insignificant except in high income-size cells.[15] Table 6 shows the relevant interactions for the column 2 specification. For the column 1 specification not shown, the only significant interactive coefficients are for income cell 5 and size cells 4 and 5 (coefficients (standard errors) of -0.204 (0.086) and -0.180 (0.074) respectively). Under the column 2 specification, there is also evidence of significant effects in large localities in income cell 4 as well, as shown in Table 6. The interactive effects are large, strongly diminishing the scale and income effects. Moving from the lowest income and size quintiles to the highest, ignoring interactions, raises servicing by 0.30; the interaction reduces that increase by about 0.20 to 0.10 for income cell 5 and sizes 4 and 5.

For other covariates in column 2, Table 5, preferences for more egalitarian policies, represented by the share of voting in the locality in favor of anti-military political parties in the 1982 elections for representatives to the *national* legislature, reveal a strong positive effect on servicing. The result also foreshadows the democratic era policies of more universal servicing. Higher own locality density also appears to increase servicing, given both cost and externality issues; a one standard deviation increase in density increases servicing by 0.05. Finally, a larger share of overall servicing in 1970 is associated with an increase in small house servicing in 1991, although the effect under urban area fixed effects is weak.

In column 3 of Table 5, we present an OLS version of column 2 adding one metropolitan area control: number of urban households. We present these results because later when we turn to examining strategic interactions we cannot include urban area fixed effects. OLS results are very similar to the fixed effect ones. Coefficients on the income-scale terms are larger, although the pattern of interactions is the same.

5.2 Robustness and Counterfactuals to the Base Case

While the vast majority of our localities are stand-alone municipalities, some are combinations of these political units. Of those localities that are multi-municipality combinations, about half have over 85% of their urban population in one dominant municipality. We reestimate the basic models in Table 5 dropping localities that are not stand-alone municipalities and dropping localities where the dominant municipality has less than 85% of the urban locality population. We also reestimate the basic models dropping all localities in the two largest urban areas, Rio de Janeiro and Sao Paulo. In all 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. For the first, we would not expect the same demand and exclusionary motivations to operate in the servicing of

[15] For income cell 2 and size cell 3, the coefficient (standard error) for the column 2 specification is -0.1523 (0.0638), while for the column 1 specification it is -0.1446 (0.0742).

large houses, which do not cater to low-income migrants, if we presume that localities have less or no desire to deflect the rich. For the second, with democratization, the national government embarked on a wide-spread policy to upgrade urban slums and their servicing. Thus by 2000, we should see less or no evidence of localities withholding service 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, based on the Table 5 column 2 (or column 1) formulation of Table 5. Although not in the table, 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 expected effects for servicing of migrant housing by existing residents that we found in Table 5 disappear in these two situations.

5.3 Strategic Interactions

We examined the existence of strategic interactions without fixed effects for the column 1 and 2 formulations of Table 5. Results on these interactions are the same for both formulations as those for the column 3 formulation of Table 5 which adds in a control for urban area size. It is on this last one that we report in Table 8. The coefficient of interest is the estimate of γ in equations (12)-(15). In column 3 for OLS estimation of equation (12), γ is positive and significant (even more so under equal weighting in the **W** matrix), but, as explained earlier, the estimate is biased upwards. Note in column 3 we try to minimize this bias by using lagged $\sum_{i\neq 1,\, i\in j} w_{1i}b_i$. However, MLE estimates of equation (15) yield small and insignificant γ's (using contemporaneous *b*'s and lagged *w*'s). There is no evidence in this formulation of strategic interactions, although the formulation as noted earlier has limitations (see also Conley's 2008 critique). Note that for the spatial correlation matrix $\mathbf{M}$, as expected given geographic contiguity, spatial correlation as measured by the ψ coefficient is positive and significant.

We also looked at other political possibilities. In particular, does the dominant locality of an urban area affect policies of other localities? We tried a simple formulation where in non-dominant localities, income and size of the dominant locality also affect service levels. Coefficients on that possibility are completely insignificant. We also tried a shock model, where between 1980 and 1991, each urban area gets an employment shock. In urban economics, the standard shock for a city is the national growth (excluding the own city) in each roughly two-digit industry for 1980-1991 multiplied by the city's base share of employment in each sector. So a city that has a big base in a nationally growing industry gets a positive shock. We constructed a shock measure for each urban area based on 27 categories of service sectors and manufacturing industries. The idea is to interact these exogenous shocks with income and size covariates for the own and/or dominant locality. If a locality is in an urban area which faces strong growth pressure, does it react by further lowering servicing to forestall increased migration pressure? Again, we found no consistent effects associated with shocks, and the base case results hold. No effects mean that own locality or dominant locality income and size terms interacted with shocks were insignificant. Again, reacting to urban area shocks requires a level of sophistication that the political system may not be able to accommodate. In summary, we think localities set their service levels for small houses occupied by migrants based just on their own base period characteristics.

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 market. Beyond enacting legal housing development restrictions, localities can withhold public infrastructure services to the informal housing 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 urban areas in Brazil between 1980 and 2000. Migrants move to these urban 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-educated. We focus on the provision of public water connections to small (1-3 room) houses in which low-income and low-educated 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 in reducing growth. Not only does withholding service from small houses reduce the growth rate of low-education households, it also reduces the growth rate of high-education households. We attribute this reduced growth effect on high-education households to externalities of living near unserviced areas.

We then estimate the determinants of 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. We find no evidence of strategic interactions among localities within an urban area in setting services.

References

Ades, A. F., & Glaeser, E. L. (1995). Trade and Circuses: Explaining Urban Giants. *Quarterly Journal of Economics , 110*, 195-227.

Au, C.-C., & Henderson, J. V. (2006). Are Chinese Cities Too Small? *Review of Economic Studies , 73*, 549-576.

Avila, P. C. (2006). *Land Use Regulations in Brazil: Impacts on Urban Markets and Access of Low-Income People to Land and Housing.* Mimeograph.

Besley, T., & Case, A. (1995). Incumbent Behavior: Vote Seeking, Tax Setting, and Yardstick Competition. *American Economic Review , 85*, 25-45.

Biderman, C. (2007). *Regulation and Informal Settlements in Brazil: A Quasi-Experimental Approach.* Mimeograph.

Black, D., & Henderson, J. V. (1999). A Theory of Urban Growth. *Journal of Political Economy , 107*, 252-284.

Brueckner, J. K. Welfare Reform and the Race to the Bottom: Theory and Evidence. *Southern Economic Journal , 66*, 505-525.

Cai, F. (2006). *Floating Population: Urbanization with Chinese Characteristics.* Mimeograph.

Case, A., Rosen, H., & Hines, J. (1993). Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States. *Journal of Public Economics , 52*, 285-307.

Conley, T. (2008). Spatial Econometrics. In S. Durlauf, & L. Blume, *The New Palgrave Dictionary of Economics, 2nd Edition.* Palgrave Macmillan.

Da Mata, D., Deichmann, U., Henderson, J. V., Lall, S., & Wang, H. (2007). Determinants of City Growth in Brazil. *Journal of Urban Economics , 62*, 252-272.

Da Mata, D., Deichmann, U., Henderson, J. V., Lall, S., & Wang, H. (2005). *Examining Growth Patterns of Brazilian Cities.* Mimeograph.

Davis, J., & Henderson, J. V. (2003). Evidence on the Political Economy of the Urbanization Process. *Journal of Urban Economics , 53*, 98-125.

Dowall, D. E. (2006). *Brazil's Urban Land and Housing Markets: How Well Are They Working?* Mimeograph.

Epple, D., & Nechyba, T. (2004). Fiscal Decentralization. In J. V. Henderson, & J.-F. Thisse, *Handbook of Regional and Urban Economics, Vol 4.* North Holland.

Fredriksson, P., & Millimet, D. (2002). Strategic Interaction and the Determinants of Environmental Policy across U.S. States. *Journal of Urban Economics , 51*, 101-122.

Gabaix, X. (1999). Zipf's Law for Cities: An Explanation. *Quarterly Journal of Economics , 113*, 739-797.

Glaeser, E., Scheinkman, J., & Shleifer, A. (1995). Economic Growth in a Cross Section of Cities. *Journal of Monetary Economics , 36*, 117-143.

Gyourko, J., Mayer, C., & Sinai, T. (2006). *Superstar Cities.* Mimeograph.

Hayashi, M., & Boadway, R. (2001). An Empirical Analysis of Intergovernmental Tax Interaction: The Case of Business Income Taxes in Canada. *Canadian Journal of Economics , 34*, 481-503.

Henderson, J. V., & Kuncoro, A. (1996). Industrial Centralization in Indonesia. *World Bank Economic Review , 10*, 513-540.

IUSS Working Group WRB. (2006). *World Reference Base for Soil Resources.* Rome: FAO.

Jefferson, G., & Singh, I. (1999). *Enterprise Reform in China: Ownership Transition and Performance.* New York: Oxford University Press.

Markstrom, S., Niswonger, R., Regan, R., Prudic, D., & Barlow, P. (2008). *GSFLOW-Coupled Ground-water and Surface-water FLOW model based on the integration of the Precipitation-Runoff Modeling System (PRMS) and the Modular Ground-Water Flow Model (MODFLOW-2005).* U.S. Department of the Interior.

Scheper-Hughes, N. (1993). *Death Without Weeping: The Violence of Everyday Life in Brazil.* Berkeley: University of California Press.

Wildasin, D. (1991). Income Redistribution in a Common Labor Market. *American Economic Review , 81*, 757-774.

Table 1
Share of Urban Households with Connection to Public Water Services

	Percent with central water connection (locality average)	Percent with central water connection (weighted by number of HHs in locality)	Number of localities with water	Total number of localities
Average for sample of 327 localities in 54 multi-locality urban 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	82	326	327
Large Houses	88	94	326	327
Migrants in poorest 20%	74	81	326	327
Non-migrants in poorest 20%	78	87	326	327
Average for localities in multi-locality urban 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 urban 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

Note: We consider localities that have less than 10% of houses served 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 its number of urban households in each decade. When the weighted share of households serviced is larger than the unweighted share, this implies that the mean level of servicing is higher in larger localities.

Table 2
Servicing of Houses by House Size

	1980	1991	2000
A. Average for 327 localities at least 50% urbanized by 1991			
Small Houses	48	74	84
Large Houses	74	88	91
Difference in Servicing	26	14	7
B. Average for 185 localities at least 50% urbanized by 1970			
Small Houses	58	81	88
Large Houses	83	93	94
Difference in Servicing	25	12	6
C. Average for 142 localities less than 50% urbanized by 1970			
Small Houses	35	65	80
Large Houses	62	82	87
Difference in Servicing	27	17	7

Note: Includes only urban households of localities in multi-locality urban areas. Small houses are those with 1-2 total rooms in 1980, covering the bottom 14.3% of the house-size distribution, and 1-3 rooms in 1991 and 2000, covering 16.5% of houses. Large houses are those with 6-7 rooms in 1980, covering 21.5 % of houses (below the top 11% by size), and 7-9 rooms in 2000, covering 19.7% of houses (below the top 6% by size).

Table 3
Growth Rate of Urban Households, 1991-2000

Dependent Variable: Growth Rate of Urban Households 1991-2000				
	(1)	(2)	(3)	(4)
	OLS	OLS	2SLS	2SLS
Share small houses with water, 1991	-0.174*	-0.159**	0.740*	0.744*
	(0.089)	(0.076)	(0.423)	(0.427)
Ln(# urban HHs, 1991)	-0.026***	0.014	-0.112***	-0.115*
	(0.009)	(0.024)	(0.030)	(0.063)
Ln (land area)	0.006	-0.012	0.060***	0.063*
	(0.010)	(0.016)	(0.019)	(0.035)
Average education in locality, 1991		-0.041***		0.000
		(0.015)		(0.049)
Share HHs rural in locality, 1991		0.260		-0.052
		(0.228)		(0.501)
Urban Area Fixed Effects	Yes	Yes	Yes	Yes
N [54 Urban Areas]	327	327	327	327
Sargan stat. p-value			0.94	0.87
Min. 1st stage partial F			13.2	13.2
Anderson LR p-value			0.11	0.06

Note: Instruments are: locality adult illiteracy rate in 1970, adult illiteracy rate in the rest of the urban 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 urban 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 urban area in 1970 interacted with the log of the distance to Sao Paulo, the log number of households in the rest of the urban area in 1970, the share of households that are rural in the rest of the urban 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.
Robust standard errors clustered at the urban area level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4
Growth Rate of Households by Education Level, 1991-2000

Dependent Variable: Growth Rate of Households by Education Level, 1991-2000 [ln (# xx educ HHs)(t) - ln (# xx educ HHs) (t-1)]	Low Education Households		High Education Households	
	(1)	(2)	(3)	(4)
	OLS	2SLS	OLS	2SLS
Share small houses with water, 1991	-0.034	0.720***	-0.277***	0.738
	(0.096)	(0.250)	(0.099)	(0.632)
Ln(density, 1991)	0.011	-0.025	-0.004	-0.099*
	(0.013)	(0.030)	(0.012)	(0.059)
Ln (#low educ. HHs, 1991)	-0.026***	-0.058***		
	(0.010)	(0.018)		
Ln (#high educ. HHs, 1991)			-0.019*	-0.036
			(0.011)	(0.027)
Urban Area Fixed Effects	Yes	Yes	Yes	Yes
N [54 Urban Areas]	327	327	327	327
Sargan stat. p-value		0.47		0.76
Min. 1st stage partial F		13.7		13.7
Anderson LR p-value		0.10		0.10

Note: Instruments are: locality adult illiteracy rate in 1970, adult illiteracy rate in the rest of the urban 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 urban 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 urban area in 1970 interacted with the log of the distance to Sao Paulo, the log number of households in the rest of the urban area in 1970, the share of households that are rural in the rest of the urban 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.
Robust standard errors clustered at the urban area level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5
Locality Decisions for Service Provision

Dependent variable: Share of small houses with water connection in own locality in 1991.
(1980 covariates unless listed otherwise)

	(1) UA fixed effects	(2) UA fixed effects	(3) OLS
Ln(median HH income)	0.485**	0.378**	0.521**
	(0.133)	(0.114)	(0.159)
Ln (# urban HHs)	0.376**	0.270**	0.309*
	(0.140)	(0.111)	(0.172)
Ln (# urban HHs)* Ln(median HH income)	-0.031**	-0.029**	-0.035**
	(0.014)	(0.011)	(0.018)
Share anti-military vote 1982 national elections		0.187**	0.185
		(0.093)	(0.116)
Ln (density)		0.033**	0.034**
		(0.114)	(0.012)
Share all urban houses with water, 1970		0.061*	0.142**
		(0.038)	(0.044)
Ln (# urban HH’s in urban area)			-0.021**
			(0.009)
N [50 urban areas]	276	276	276
R^2 [within]	[0.23]	[0.30]	0.44

Robust standard errors clustered at the urban area level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 6
Income-Size Interactions in Setting of Service Levels

Dependent variable: Share of small houses with water connection in own locality in 1991				
	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Ln(median HH income)	0.133**	0.140**	0.160**	0.191**
	(0.049)	(0.053)	(0.075)	(0.069)
Ln (# urban HHs)	0.041	0.080	0.120*	0.111
	(0.062)	(0.056)	(0.063)	(0.069)
		Relevant interactions		
4th income*4th size quintile		-0.141**		
		(0.070)		
4th income*5h size quintile		-0.100		
		(0.089)		
5th income*4th size quintile		-0.225**		
		(0.069)		
5th income*5th size quintile		-0.194**		
		(0.064)		
Includes urban area fixed effects and other covariates from column 2 of Table 5		Yes		
N [50 urban areas]		276		
R^2 within		0.35		

Robust standard errors clustered at the urban area level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 7
Counterfactuals

Servicing of large houses in the non-democratic are and of small houses in the democratic era.
(Covariates lagged by one period unless listed otherwise)

	(1) Service to large houses in 1991	(2) Service to small houses 2000
Ln(median HH income)	0.102	0.065
	(0.097)	(0.147)
Ln (# urban HHs)	0.072	0.026
	(0.087)	(0.101)
Ln (# urban HHs)* Ln(median HH income)	-0.0070	-0.0047
	(0.0087)	(0.013)
Share anti-military vote 1982 national elections	0.039	0.152*
	(0.077)	(0.087)
Ln (density)	0.030**	0.038**
	(0.0084)	(0.011)
Share all urban houses with water, 1970	0.050	0.090**
	(0.035)	(0.043)
Urban area fixed effects	Yes	Yes
N [50 urban areas]	276	276
R^2 within	0.29	0.23

Robust standard errors clustered at the urban area level.
* significant at 10%; ** significant at 5%; *** significant at 1%

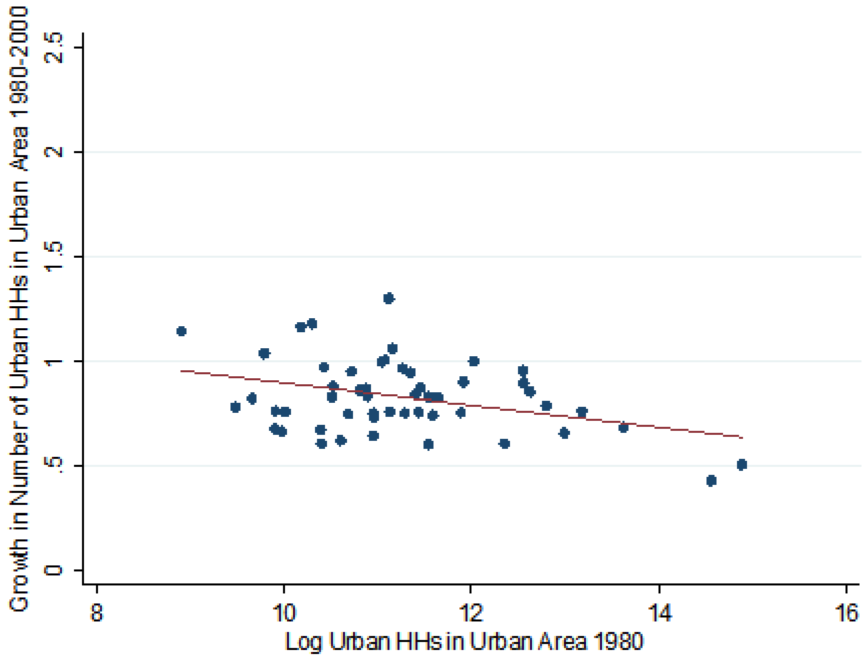
Table 8
Testing for Strategic Interactions

Dependent variable: share of small houses with water connection in own locality in 1991	(1) MLE (Inverse Density Weights)	(2) MLE (Equal Weights)	(3) OLS (Inverse Density Weights)
Weighted average of other localities' in UA share of servicing of small houses	0.010	0.113	0.289**
	(0.114)	(0.128)	(0.053)
Spatial correlation	0.475**	0.400**	n.a.
	(0.084)	(0.108)	
Extended controls	Yes	Yes	Yes
N [50 urban areas]	276	276	276
R^2	0.58	0.58	0.44

Note: Columns 1 and 2 use 1991 service levels in other localities. Column 3 uses 1980 to try to mitigate endogeneity issues.
Robust standard errors clustered at the urban area level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Figure 1
Urban Area and Locality Growth in Urban Households, 1980-2000

a) Urban Areas [slope coefficient (standard error) of -0.0530 (0.0184), R^2 =0.14; dropping Sao Paulo and Rio de Janeiro: -0.0287 (0.0219), R^2 =0.03]



b) Localities [slope coefficient (standard error) of -0.106 (0.012), R^2 =0.18]

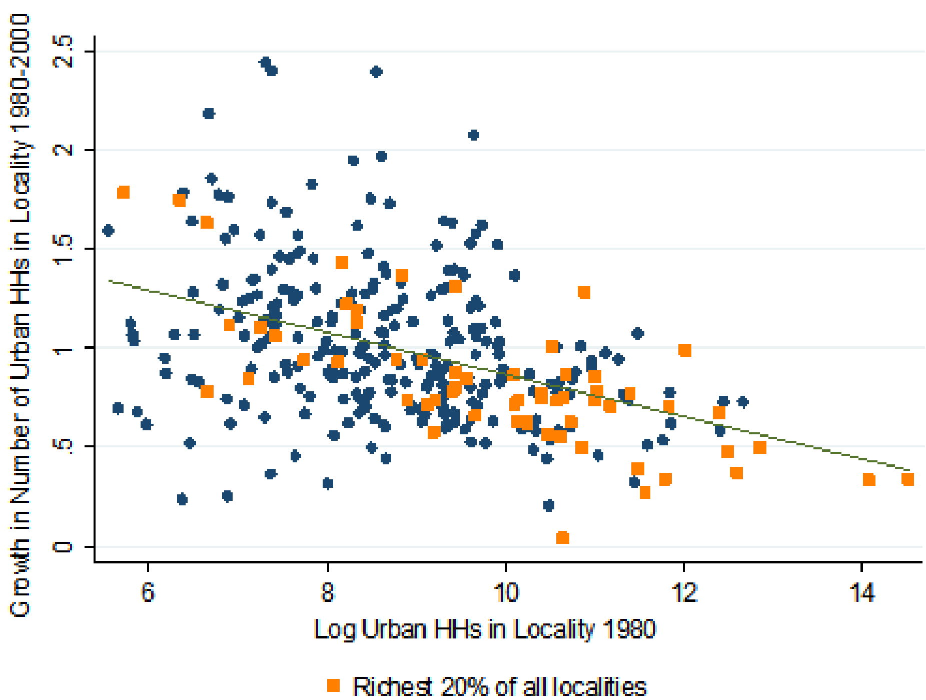
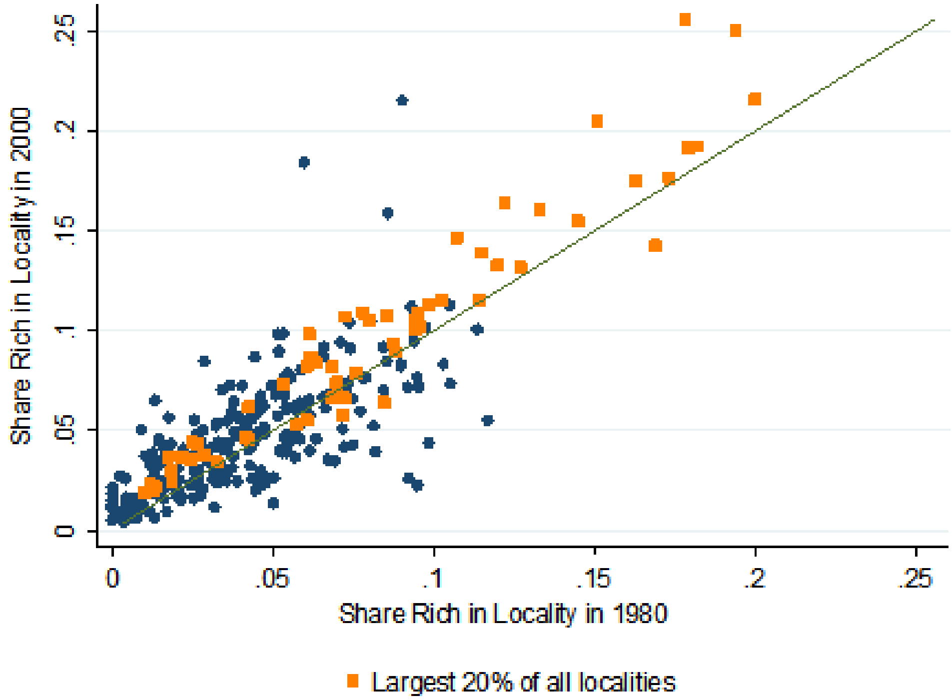


Figure 2
Share Rich in Localities in 1980 versus 2000



Regression (standard errors in parentheses): Share Rich 2000=0.0124 (0.0021) + 0.8023 (0.0447)*Share Rich in 1980 + 0.2331(0.0410)*Share Rich in 1980* dummy for largest 20% of localities. R^2=0.77.

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. 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. The 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 has 56 and Rio de Janeiro 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.

For services, we do a full examination of all types and forms of services on columns 1 and 3 of Table A1 and then a reduced form in columns 2 and 4 where we use the typical summary measures: central water connection and full service (any electricity, central sewer, and connection to central piped water). Table A1 shows the basic results.

In Table A1, the reported coefficients reflect the percent by which rents rise. From columns 1 and 3, it is clear that in both Sao Paulo and Rio de Janeiro, there is a high premium on having central water piped into the house, substantially more than well water piped into the house, presumably reflecting the greater reliability of supply. Electricity garners a very large premium, even more so if it is metered (legal), indicating both reliable supply and higher (amperage) effective service. Public garbage collection has modest or no impacts. Central sewer is much more valued than septic systems, especially in Rio de Janeiro. Septic systems only raise premiums modestly above having no service, presumably reflecting the failure of septic systems in these dense localities. 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

	Dependent variable: Ln(Rent), 1980					
	Sao Paulo			Rio de Janeiro		
	(1)	(2)	Mean	(3)	(4)	Mean
Water: other: inside plumbing	0.121**		0.0017	0.161**		0.0038
	(0.030)			(0.040)		
Water: well, inside plumbing	0.185**		0.023	0.201**		0.0067
	(0.011)			(0.032)		
Water: central connection, exterior	0.121**		0.055	0.229**		0.043
	(0.009)			(0.019)		
Water: central connection, interior	0.285**		0.897	0.366**		0.93
	(0.020)			(0.018)		
Septic system	0.022**		0.308	0.028**		0.049
	(0.005)			(0.014)		
Central sewer	0.094**		0.652	0.177**		0.916
	(0.006)			(0.012)		
Sanitation: for more than one house	-0.0029		0.191	-0.057**		0.091
	(0.018)			(0.030)		
Sanitation: own house collection	0.060**		0.804	-0.034		0.903
	(0.018)			(0.031)		
Electricity: no meter	0.077**		0.284	0.192**		0.162
	(0.013)			(0.033)		
Electricity: meter	0.189**		0.708	0.342**		0.833
	(0.013)			(0.034)		
Central water inside house		0.181**			0.214**	
		(0.007)			(0.015)	
Full service (central water inside house, electricity, central sewer)		0.195**			0.361**	
		(0.004)			(0.007)	
Controls: House Characteristics, District Fixed Effects	Yes	Yes		Yes	Yes	
N	196,149	200,067		105,135	108,492	
Districts	56	56		24	24	
R^2	0.65	0.50		0.58	0.54	

Appendix B:

Table B1
Summary Statistics

For Tables 3 and 4: N=327, 54 Urban Areas	Mean	Std Dev	Min	Max
Share small houses with water 1991	0.74	0.21	0.02	1.00
Growth rate: # HHs 1991-2000	0.40	0.17	-0.07	1.05
Growth rate: # low-education HHs 1991-2000	0.54	0.20	-0.16	1.20
Growth rate: # high-education HHs 1991-2000	0.33	0.22	-0.25	1.29
Ln(# urban HHs in locality) 1991	9.50	1.46	6.10	14.73
Ln(# low-education urban HHs in locality) 1991	8.51	1.34	5.43	13.32
Ln(# high-education urban HHs in locality) 1991	9.00	1.56	5.39	14.45
Ln(density) 1991	5.11	1.73	0.22	9.42
Ln(land area) 1991	5.98	1.36	2.80	12.23
Average schooling in locality 1991	4.71	1.10	2.06	8.84
Share HHs rural in locality 1991	0.14	0.13	0.00	0.51

For Tables 5, 6, 7, and 8: N=276, 50 Urban Areas	Mean	Std Dev	Min	Max
Share small houses with water 2000	0.86	0.15	0.10	1.00
Share small houses with water 1991	0.77	0.19	0.02	1.00
Share small houses with water 1980	0.53	0.26	0.00	1.00
Share large houses with water 1991	0.90	0.13	0.14	1.00
Ln(# urban HHs in locality) 1991	9.81	1.33	6.34	14.73
Ln(# urban HHs in locality) 1980	9.25	1.41	6.30	14.52
Ln (median HH income) 1991	7.10	0.38	6.00	7.94
Ln (median HH income) 1980	9.52	0.39	8.41	10.44
Share of locality voting for anti-military parties 1982	0.59	0.15	0.03	0.94
Ln(density)80	5.10	1.64	0.64	9.36

Table B2
First Stage Regressions

Dependent Variable:	Share small houses, central water 1991	Ln(# urban HHs) 1991	Average education 1991	Share HHs rural 1991	Ln (#low educ. HHs) 1991	Ln (#high educ. HHs) 1991	Ln(density) 1991
Ln(land) 91	-0.021*	0.349***	0.079**	0.021***			
	(0.011)	(0.059)	(0.030)	(0.008)			
Illiteracy 70	-0.007***	-0.074***	-0.077***	0.004***	-0.044***	-0.072***	-0.089***
	(0.001)	(0.008)	(0.007)	(0.001)	(0.007)	(0.008)	(0.010)
Illiteracy rest UA 70	-0.027	-0.007	0.163	0.017	-0.109	-0.069	0.090
	(0.017)	(0.132)	(0.120)	(0.012)	(0.109)	(0.123)	(0.168)
Share anti-military vote, nat.	0.163	2.068***	0.046	-0.184***	1.949***	2.032***	2.040***
legislative elections 1982	(0.101)	(0.455)	(0.272)	(0.061)	(0.374)	(0.427)	(0.615)
Manu-service employ ratio, rest UA	-0.128	2.724**	-0.284	-0.056	3.581***	3.329***	1.476
	(0.304)	(1.101)	(1.345)	(0.080)	(1.213)	(1.233)	(1.063)
Manu-service employ ratio, rest	0.013	-0.446**	0.136	0.015	-0.613**	-0.538**	-0.230
UA*ln(dist to SP)	(0.056)	(0.206)	(0.243)	(0.015)	(0.232)	(0.235)	(0.197)
Illiteracy rest UA 70* ln(dist to SP)	0.003	-0.012	-0.029*	-0.001	-0.001	-0.009	-0.017
	(0.002)	(0.017)	(0.016)	(0.002)	(0.014)	(0.016)	(0.021)
Ln(# urban HHs, rest UA) 70	-0.091**	-0.946**	-0.282**	0.033	-1.552***	-1.702***	0.345
	(0.038)	(0.388)	(0.136)	(0.029)	(0.363)	(0.378)	(0.348)
Share HHs rural in rest UA 70	-0.042	1.719	2.329**	-0.370**	0.210	0.249	3.905*
	(0.169)	(1.753)	(0.993)	(0.154)	(1.448)	(1.535)	(2.082)
% geology porous	-0.489**	-0.634	-0.712	-0.245	-0.541	-1.227	-0.244
	(0.239)	(1.437)	(0.969)	(0.209)	(1.584)	(1.473)	(1.582)
Insolation mean	-0.085	0.151	0.326	-0.224	0.705	0.280	-0.633
	(0.207)	(1.008)	(0.736)	(0.164)	(1.109)	(1.193)	(1.492)
Insolation standard deviation	0.165	6.683	4.108***	-1.180**	6.486	6.888	4.869
	(0.459)	(5.512)	(1.237)	(0.534)	(4.882)	(4.781)	(6.728)
% geology porous*insolation mean	0.332***	0.249	0.368	0.143	0.175	0.613	0.057
	(0.109)	(0.710)	(0.536)	(0.117)	(0.752)	(0.716)	(0.867)
% geology porous*insolation	-0.593***	-1.298	-1.045	0.066	-1.163	-1.686	-1.015
standard deviation	(0.191)	(1.191)	(1.027)	(0.174)	(1.381)	(1.306)	(1.617)
N [54 Urban Areas]	327	327	327	327	327	327	327
F-statistic	15	125	76	16	59	103	50

Robust standard errors clustered at the urban area level.
* significant at 10%; ** significant at 5%; *** significant at 1%