

NBER WORKING PAPER SERIES

EXCLUSIONARY POLICIES IN URBAN DEVELOPMENT, HOW UNDER-SERVICING  
OF MIGRANT HOUSEHOLDS AFFECTS THE GROWTH AND COMPOSITION  
OF BRAZILIAN CITIES

Leo Feler  
J. Vernon Henderson

Working Paper 14136  
<http://www.nber.org/papers/w14136>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
June 2008

We thank Andrew Foster and seminar participants at Carnegie Mellon and Brown Universities and at the Philadelphia Federal Reserve for helpful comments and suggestions. We are grateful to Cassio Magalhaes, Flora Maravalhas, and Adam Storeygard for research assistance, to Lynn Carlson for assistance with GIS, and to Cristina Costa at the Brazilian Meteorological Institute (INMET) for providing data. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2008 by Leo Feler and J. Vernon Henderson. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Exclusionary Policies in Urban Development: How under-servicing of migrant households affects the growth and composition of Brazilian cities

Leo Feler and J. Vernon Henderson

NBER Working Paper No. 14136

June 2008

JEL No. D7,H7,J6,O15,O54,R5

### **ABSTRACT**

Localities in developed countries often restrict construction and population growth through regulations governing land usage, lot sizes, building heights, and frontage requirements. In developing countries, such policies are less effective because of the existence of unregulated, informal housing markets. Cities in developing countries that seek to limit in-migration must also discourage entry into informal housing by providing low levels of public services to this sector. In this paper, we analyze the causes of slums, using data from Brazilian urban areas. We develop a model of the decisions that localities make to affect in-migration and find evidence that localities act strategically. Richer and larger localities in an urban area reduce provision of water and sewerage connections to the smaller houses in which poorer migrants would live to discourage the in-migration of these poorer migrants and deflect them to other localities. We also find that under-servicing smaller houses reduces the population growth rate of localities. Not only does it reduce the in-migration of low-educated households, it seems that, because of negative externalities, such under-servicing also reduces the growth rate of higher-educated households.

Leo Feler

Department of Economics

Box B

Brown University

Providence, RI 02912

leo\_feler@brown.edu

J. Vernon Henderson

Department of Economics

Box B

Brown University

Providence, RI 02912

and NBER

j\_henderson@brown.edu

# 1 Introduction

In developing countries during periods of rapid urbanization, urban areas often house significant portions of their populations in informal housing sectors. From the 1960s through the 1990s, this was illustrated by the development of *favelas* and *loteamentos* in Brazil and similar types of settlements in other Latin American countries. Today, development of such settlements is played out in the slums of Sub-Saharan Africa and in the “urban villages” of Beijing and other Chinese urban areas. Informal housing sectors are usually characterized by varying degrees of insecurity of tenure, but, perhaps more critically in some contexts, they tend to be cut off from basic urban services such as central water and sewerage. This makes living conditions unpleasant, unhealthy, and expensive.

While unserviced settlements may reflect a failure in governance or the low incomes of residents, we hypothesize that there is a strategic element involved. Provision of bad living conditions for migrants is a way for existing residents to discourage in-migration to a locality, particularly of low-income migrants. Indeed, in China today, this strategic component is articulated as an explicit policy currently adopted by the largest urban areas (Cai (2006)). Forcing the vast majority of migrants into poorly serviced “informal sector” settlements is viewed as a key element in restraining rural in-migration to these urban areas overall.<sup>1</sup> In Brazil, there has never existed such an explicit articulation of exclusionary housing policy as in China. The localities or districts of Brazilian urban areas make much more decentralized policy decisions that influence in-migration to their specific localities. Nevertheless, corresponding forces seem to be at work in Brazil. We examine how certain localities within an urban area in Brazil may act to induce migrants to locate in other localities of the same urban area—providing an element of a “race to the bottom” in terms of servicing migrants. This limits in-migration to such localities within the urban area; but of course such policies may also inhibit overall migration to the urban area.

The development of unserviced informal housing sectors has immediate effects beyond restraining migration: inequality in living conditions and development of unhealthy neighborhoods with high negative externalities. Poor servicing may affect the location decisions of low-skilled migrants to a locality; but the resulting negative externalities may also affect the supply of high-skilled labor. These policies have implications for the future as countries develop and undertake investments to make cities more “livable”, which will involve catch-up investments. Building water and sewer infrastructure long after the development

---

<sup>1</sup>These areas are informal in the sense that (1) they are under rural governance, even when situated well within the city limits; (2) the city is not responsible for servicing them; and (3) the migrants do not have rights of local citizens (rural or urban).

of dense neighborhoods can be very costly, requiring extensive spatial reconstruction and reconfiguration of neighborhoods.

To better understand this process, we study Brazil from 1980-2000. While Brazil is urbanized today, and the north to south movement of people has diminished, local resistance to in-migration of low-skilled migrants was an important issue in the 1980s, with effects on urban growth persisting through the 1990s. By studying Brazil in this time period, we hope to better understand the forces at work in China and in other urbanizing countries today. For Brazil, we examine (1) public infrastructure investment behavior in localities in the 1980s and (2) the impact of such policies on subsequent locality population growth and social composition in the 1990s. We find evidence consistent with strategic behavior and find that such policies affected population growth rates of localities.

There is an extensive literature on exclusionary policies of local jurisdictions (e.g., the Tiebout literature, as reviewed in Epple and Nechyba (2004)). Building on this literature, more recently Gyourko, Mayer, and Sinai (2006) analyze the development in the United States of what they call “superstar” cities. These are cities typically favored with excellent natural amenities, where population growth has slowed, the share of the population from higher income groups is high and increasing over time, 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, might dissipate the advantages of the high natural amenities. If high-income consumers have a higher willingness-to-pay per unit of housing for these amenities, they will outbid lower-income consumers to live in superstar locations. If the national population is growing over time, superstar cities will become increasingly relatively richer as they draw from an increasingly smaller upper tail of the income distribution.

The reasons for exclusion in developing countries have the related stratification and amenity elements. But there are two key differences. The first involves motivation and setting; the second, a more substantive analytical issue. For the first, certain urban areas may not be favored so much by natural amenities but rather by policy initiatives of national governments in terms of capital market allocations for industry, provision of public services for incumbent residents, licensing for export, foreign direct investment, imports, and government investment in state capitalism. The literature makes this point generally (Ades and Glaeser (1995) and Davis and Henderson (2003)) and then with examples from Indonesia and China, (Henderson and Kuncoro (1996), Jefferson and Singhe (1999), and Au and Henderson (2006)), as well as Brazil. This favoritism attracts migrants seeking job opportunities. If in-migration is unfettered, in the end, such favored urban areas become

“over-populated”, in the sense that migration only ceases when the increased congestion, living costs, and diminished quality of life from over-population lead to dissipation of the benefits of national government favoritism.

Not surprisingly, incumbent residents of localities within an urban area may seek to restrict in-migration to their locality of at least low-income people, to halt this dissipation process within their locality. Such a situation may represent conflicting interests at the local, state, and national government levels. While localities may want to serve the interests of incumbent residents by limiting in-migration, national and state government officials may have a different political agenda, favoring certain urban areas with subsidized investment and public services such as schooling, implicitly encouraging overall migration to these urban areas.

Given the desire to exclude, the second difference in application to developing countries involves the operation and nature of restrictions. In developed countries, exclusion occurs through formal market housing restrictions—development fees, zoning, and other policy levers limiting housing development and population density. In many developing countries, formal housing sector restrictions, rather than halting locality growth as in the United States, lead to development of an informal sector where such restrictions are ignored, which, unlike in the United States, is “tolerated”. By tolerated we mean it is not politically feasible to halt development of informal settlements and/or institutions are sufficiently weak, making enforcement of a ban on informal settlements impossible. Given formal sector restrictions are not enough to halt growth, exclusion defaults into informal sector policies, which may retard the rate of growth.

Brazil has two types of informal sector housing markets. First are *favelas*, which were historically created by land 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 other parties than the occupier. Second are *loteamentos*, where developments do not meet zoning regulations, but are built on legally acquired land. However, after development, owners cannot obtain land title because the 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.

*Loteamentos* are a more recent development, 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 low- and low-middle income families. Since then, individual localities have

imposed even stricter minimum lot size requirements. As urban areas expanded after 1979, a substantial part of the increased housing demand was met by suburban developments that violated the national zoning law.<sup>2</sup>

A key aspect is that until the late 1980s and democratization, it was in principle “illegal” for localities to provide public infrastructure such as central sewer and water connections in either type of settlement. This gave an opportunistic excuse to deny or limit such provision. While localities cannot effectively halt informal sector development, they can contain it by denying basic public infrastructure services to such neighborhoods, with a twofold impact for migrants—poor living conditions and the need to substitute expensive private alternatives to public provision of basic services.

Section 2 of the paper discusses data and trends towards exclusion. Section 3 discusses a conceptual framework to inform econometric specifications. Section 4 analyzes whether localities in a metropolitan area seem to interact strategically to exclude migrants and what types of localities are more likely to “under-provide” public infrastructure to migrants. Section 5 estimates the impact of exclusion on locality population growth and population composition. Section 6 concludes.

## 2 Urbanization and Public Infrastructure in Brazil

This section provides background information on Brazil relevant to our analysis, both theoretical and econometric. 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 informality and then turn to the issue of how, in the data, to represent policy initiatives which are based on exclusionary considerations.

The paper focuses on the post 1980 time period. We look at public service provision during the 1980 time period, which starts off as non-democratic. Full democratization at the national level occurs in 1988, with democratic reforms in subsequent years.<sup>3</sup> One set of reforms removed restrictions on localities providing infrastructure to the informal sector; another encouraged the regularization of informal housing sectors and the upgrading of services. Our working presumption is that during the 1980s, exclusionary behavior by localities in terms of service provision was possible, even though most of our localities had elected and not military appointed mayors. Elitist dominated cities could legitimately deny

---

<sup>2</sup>Indeed although suburban residents are 37.4% of all urban households in major localities, 43.2% of homes reporting no land title are in suburban areas in 2000.

<sup>3</sup>1988 dates the new constitution. The government moved from military to civilian control in 1985.

services to the informal sector. However, by the 1990s, following reforms, such strategic behavior was more problematic.

After looking at service provision in the 1980s, we will examine subsequent locality population composition and growth in the democratic 1990s. 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 transport and telecommunications, as well as agriculture developments in the north (da Mata, Deichmann, Henderson, Lall, and Wang (2005)). By the 1990s, these adjustments are largely complete. This change in urbanization and industrialization patterns reflects a change in the underlying drivers of city growth from the 1970s to the 1990s. This change in drivers will provide a basis for one aspect of the instrumental variables approach in Section 5.

## 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 regulation, we do have retrospective information. A census of local governments conducted in 1999 and in 2005 [IBGE, *Perfil dos Municípios Brasileiros*, 1999 and 2005] indicates whether cities had passed a minimum lot size zoning law in excess of the national standard of 125 square meters by 1999.<sup>4</sup>

Local governments in Brazil are municipalities (“municípios”), units equivalent to United States counties, with larger urban areas consisting of multiple municipalities. Our unit of analysis, in principle, is the municipality, as well as 123 urban areas (defined as of 1991) consisting of either a single municipality or of a collection of municipalities (59 of the 123 cases). Over time, analysis must account for the fact that some initial municipalities split into more municipalities, and there were some recombinations as well.<sup>5</sup> To facilitate over-time analysis, we combine split municipalities into “common denominator” ones, which we call localities (informally) or MCAs (Minimum Comparable Areas) more formally (see da Mata, Deichmann, Henderson, Lall, and Wang (2005)). Most multi-municipality MCAs

---

<sup>4</sup>Data are available from IBGE, *Perfil dos Municípios Brasileiros*, 1999 and 2005.

<sup>5</sup>Fiscal transfers from the federal government after 1988 favored creation of new municípios (formulas included a lump sum component independent of size).

are dominated by one municipality, where that municipality contains over 85% of the urban population of the MCA.<sup>6</sup> The 123 urban areas are composed of 447 localities (MCAs), which in 1991, consisted of 659 municipalities. Urban areas are either defined as agglomerations by the Institute of Geography and Statistics (IBGE), or are single localities with over 75,000 people that are over 75% urbanized in 1991. Since what is urban changes dramatically in the 447 localities from 1970 to 2000, especially in the suburban localities, we often look at sub-samples of localities, imposing criteria such as requiring the locality to be at least 50% urbanized in a given census year.

## 2.2 Patterns: Urban Growth and Stratification

Our spatial unit of analysis is the locality of an urban area. In making this choice, we were informed by broad patterns in the data. Urban areas in Brazil experienced “parallel growth” from 1980–2000, meaning that small and large urban areas grew at about the same rate, as a number of theories predict (e.g., Black and Henderson (1999) and Gabaix (1999)) and Figure 1a demonstrates. While the dispersion of growth rates is larger for smaller urban areas in Figure 1a, there is no evidence of “convergence” or relative mean reversion for urban areas. Urban areas grow overall in parallel, with knowledge accumulation and improved education levels (da Mata, Deichmann, Henderson, Lall, and Wang (2007)). In contrast, in Figure 1b, localities within urban areas experience significant mean reversion: bigger localities, often central city ones, grow at a slower rate than smaller localities. Note Figures 1b and 2 distinguish localities in the top 10% by urban population size, center city localities, localities which are both center cities and in the top 10% by size, and all others.

Overall these differing growth patterns between urban areas and localities represent two factors. With population growth in an urban area, old localities “fill-up” and become crowded, and new localities develop. Second with economic growth, urban areas spread out and suburbs develop, fueled as well by declining commuting costs with transport 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 empirical work. We note that in Brazil, as in much of the world, the rich live predominately in the center cities, and the poor live in the suburbs. Thus, in terms of exclusion, we are thinking of central

---

<sup>6</sup>For most of our analyses we restrict our sample to those MCAs that are predominantly urban (have at least 50% of their population living in urban areas) in 1991 and are located in urban areas with other MCAs. This leaves a sample of 351 MCAs in 55 urban areas. For this sample, 258 of the 351 MCAs are comprised of just a single municipio (i.e., there were no recombinations or splits of municipios since 1970). Another 48 have at least 85% of their population in one “dominant” municipio and another 20 at least 65% of their population in a “dominant” municipio. For the remaining 25 MCAs, the most populous municipio has between 20% and 64% of the MCA population.

cities as well as larger and richer suburbs deflecting migrants and low-income residents into low-income suburbs.<sup>7</sup> These notions are reinforced by the fact that overtime richer localities have grown richer. In Figure 2, we plot the percent rich in 2000 against the percent rich in 1980. We see that larger localities have improved their share of the rich in general: their data points lie above the 45-degree line. For example, given that center cities are generally the largest locality in the urban area, the slope coefficient of 1980 share rich on 2000 share rich has a slope coefficient for central cities of 1.06, which is significantly higher than 0.94 for all localities.

## 2.3 The Informal Sector and Public Infrastructure

An issue is how to identify those living in the informal sector in Brazil (e.g., Dowall (2006); Biderman (2007)). Studies suggest large portions of the population live in the informal sector, around 30% for bigger cities, although criteria for informality differ as do estimates. We summarize the possibilities for defining the informal sector as discussed in the literature. In the census, there is a question filled out by census takers on whether people live in “irregular settlements”. Irregularity concerns whether streets are straight or crooked, houses properly numbered in a neighborhood configuration of housing, not whether houses are serviced or owners have formal title. Thus, irregularity differs from informality, and less than 5% of households are considered irregular. Economists typically prefer to define informality based on ownership rights. In the 1991 and 2000 Census, for home owners, there is also a question on whether home owners have title to their land. In 1991, about 8% of all urban households in our urban localities live in owner-occupied housing for which they do not report land title, which corresponds to about 13-14% of owner-occupiers. Again the number seems small; 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 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’s 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 (Dowall (2006)) suggests a key element is a central water connection,

---

<sup>7</sup>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.

where in 1991, about 14% of urban households overall are not connected. 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 51% of households do not have full service. For water, lack of a water connection means private alternatives must be used. Historically in many localities, especially those situated on large water tables, the private alternative was to dig a well. Today, as water tables are strained with wells running dry for portions of the year and as population density in suburban localities increases, the private alternative at the margin for a migrant is to use a public stand pipe (and haul the water for some distance), subscribe to a water truck service (*carros-pipa*), or purchase bottled or bagged water. These are difficult or expensive alternatives.

As a preliminary check that central connections are highly valued, we examine willingness-to-pay for such services, 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 main 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 center city are available, so we can use 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 additional premium for indoor plumbing) in Sao Paulo and 23% in Rio de Janeiro. 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. Breaking out 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 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).

In this paper, we focus on the notion of exclusion through lack of servicing, whether land is reported to have formal title or not and whether housing meets land-use regulations or not. But how do we represent exclusion through lack of servicing? This is tricky because in Brazil (unlike China) exclusionary policies cannot target individuals based on personal characteristics such migration status. They can only target neighborhoods where migrants are likely to live, recognizing migrants can live where they demand given their incomes and prices. We start by looking at the evolution and extent 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 servicing. We look at localities that are at least 50% urbanized by decade, as well as localities that are at least 50% urbanized in 1970 (allowing us to track service expansion within a constant sample of localities). For those places that have “no service” in 1970 and are at least 50% urbanized, we track how quickly services expand in the subsequent decades. For this table, we define localities that have less than 10% of houses served as having “no service”. This allows for error in reporting of servicing (especially in finer categories of types of households) and for situations where only one special neighborhood of a locality has service for idiosyncratic reasons.

Table 1 shows the rapid expansion in services in urban Brazil over the decades. In 1970, localities that were at least 50% urbanized provided a connection to a central water system to 51% of their urban households. By 2000, this number had 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. For full servicing, where the shortfall relative to a water connection is usually just lack of central sewer, in 1970, localities provided full service to 25% of their urban households, with this number reaching 49% by 2000. The fact that the weighted<sup>8</sup> percentage of houses with a central water connection or with full service is higher than the unweighted percentage suggests that more populous localities service a greater share of their households.

The second panel of Table 1 shows similar results for a constant sample of 258 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 to which they provide service, from a mean of 51% in 1970 to 92% by 2000. Finally, in the third panel, we explore the increase in servicing among places that were at least 50% urbanized but had “no service” in 1970. Within the course of a decade, from 1970 to 1980, these localities grow from having essentially no houses connected to a central water system (2%) to having 57% of houses connected. This suggests that localities can quite rapidly expand their central water systems, and the fact that some houses remain unserved even by 1991 and 2000 might reflect a strategic element to servicing.

The fifth column of the table shows one factor that will lead us to focus on water provision rather than full servicing when examining exclusion and its effects. Central water service generally exists in most urbanized localities; and the margin we examine is the extent of its provision within the locality. By 1990, 428 of the 429 urbanized localities have central

---

<sup>8</sup>We weight the percentage of houses serviced with a central water connection or with full service by the urban population of the locality.

water for at least 10% of the households. However, many localities have absolutely no full service, meaning no central sewerage. In 2000, 17% of our localities have no such service, and in 1991 it is 42%. For localities with no full service, it is obviously impossible to analyze exclusionary behavior in terms of denying full servicing to the informal housing sector.

There are two other reasons we focus on water connections. First, provision of a central water connection seems to generally be more of a locality decision made by municipality water authorities or in negotiation with regional authorities, while sewer provision seems more of a state-level decision. Second, the history of spatial development in a city is a key element determining central sewer connections, but less so for water. While water connections to unserved areas are easy to add using above-ground water connections, sewers require a major investment and upheaval in terms of digging up streets and even house demolition to bring service to historically unserved areas. Many neighborhoods, even richer ones, without sewer connections continue to rely on private alternatives in septic systems.

The first panel of Table 1 also explores spatial, housing tenure, and income differences in service provision for households in 1991. Suburban areas, with their low population densities and lower incomes, have poorer servicing than localities in general. In terms of housing tenure, the vast majority of households that report owning their homes are well serviced, although the best served category are renters, who live in the core parts of older cities, which are “grandfathered” with central water and sewer. Not surprisingly, the worst served are those who report that they do not own the land under their house. Similarly, those living in rent-free or ceded housing, such as employer provided or temporary squatter housing, are poorly serviced. In the last rows of the first panel, we see that low-educated households, where the household head did not complete primary school, are more poorly serviced than higher-educated households (where the household head completed at least primary school). While migrants are more poorly 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 10 years) are more poorly 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 poor in the population, potentially welcoming higher-income migrants (who may displace existing lower-income migrants) and discouraging lower-income migrants. 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 *per se*, nor migration status. What localities can do is not service the houses that most migrants and low-educated households are likely to occupy.

These tend to be the smallest houses with 1–2 total rooms (in 1970 and 1980) or 1–3 total rooms (in 1991 and 2000). In 1991, 33% of the migrants in our localities lived in 1–3 room (small) houses and only 12% lived in 7–9 room (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, 27% lived in small houses and 13% lived in large houses, compared with 17% and 19%, respectively, for higher-educated households. Targeting the smallest houses therefore appears to be an effective mechanism in discriminating against migrants and low-educated households.

In Table 2, we examine the service levels for small houses likely to be occupied by migrants and low-educated households: 1–2 rooms in 1980 covering the bottom 14.3% of the house-size distribution and 1–3 rooms in 2000 covering 16.5% of houses. We compare these with houses for upper-middle-income households: those with 6–7 rooms in 1980 covering 21.5% of households (below the top 11% by size), and 7–9 rooms in 2000 covering 19.7% of households (below the top 6%). Table 2 gives weighted averages for servicing of small versus large houses. In 1980, only 61% of small houses in our localities had a water connection while 86% of large houses had a water connection. The relative difference in full servicing is even more dramatic: 17% versus 54%. By 1991, the water gap diminishes but is still noticeable overall, and even more so in individual localities.<sup>9</sup> In Sections 3 and 4, we will use the provision of a public water connection to small houses as our basic “exclusionary” measure, representing quality of infrastructure faced by incoming residents.

## 2.5 Land-Use Regulations

In Table 3, 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 as noted earlier. In Table 3, we list what fraction of significant size localities (with over 15,000 urban residents) by 1999 had passed a minimum lot-size law in excess of 125 square meters—that is, a minimum lot-size law in excess of the national standard. In column 2, we look at the ratio of urban households in 2000 that are migrant relative to non-migrant. We compare these ratios in locally zoned localities—those with a more stringent minimum lot-size law than the national law—to those without local zoning. In column 3, we list the ratio of homeowners without title to those with title for the same groups of localities. Note more migrants relative to non-migrants and more households without title relative

---

<sup>9</sup>The unweighted averages for water in 1991 for small versus large houses are 57% and 89%, respectively.

to those with title tend to live in areas with minimum lot-sizes in excess of the national standard. This hints at the endogeneity of regulations: localities that impose stronger zoning regulations may be those subject to migration pressure. Endogeneity will be critical in the later identification of the effects of servicing on locality growth.

### 3 Conceptualizing Exclusionary Behavior

This section develops a simple model upon which we base the empirical formulations of strategic behavior, servicing levels, and population growth. Since the empirical work is focused on within urban area variation for identification of effects, we are not going to focus on the determination of urban area characteristics. For example, we assume workers in all localities in an urban area 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 (although in the empirics we experimented with implicitly allowing the choice to affect disposable incomes after commuting costs).

We formulate the basic problem much like the welfare competition literature in the United States (Wildasin (1991), Brueckner (2000)), where within a region, localities are choosing a policy tool as they face a potential influx of migrants. In our case, the policy tool 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. Localities value better services for these migrants for either altruistic or externality reasons, which is a force to increase service levels. However, they would prefer fewer migrants to their own locality, which is a force to reduce service levels, although, for economic growth reasons, they may want more migrants overall to the urban area. 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, the service levels in other localities, and the characteristics of the own locality’s base-resident population, each locality strategically chooses a level of servicing. We outline a general functional form model that captures the key aspects and then illustrate a specific version which gives simple estimating equations.

Migrants have preferences of the form

$$U = \hat{U}(x, h, g, b), \tag{1}$$

where  $x$  is the numeraire good,  $h$  is housing,  $g$  is the quantity of urban services such as water, and  $b$  is the share of migrants who are serviced, a positive externality for migrants and residents. The share of migrants served is a policy variable chosen by the locality. The endogenous number of migrants to the locality is  $L$ . All migrants are assumed to live in the informal housing sector while base-residents live in the formal sector. Based on policy decisions by the locality, some migrants will live in neighborhoods where the locality publicly provides a reasonable quality of public services at a unit cost  $c_0$  (e.g., the cost of metered water). Other migrants will live in unserviced neighborhoods, where residents must privately secure services at a higher unit cost,  $c$  (e.g., water purchased from water delivery trucks).

### 3.1 Equilibrium within the Informal Sector

Migrants residing in serviced neighborhoods have housing demand functions and quasi-indirect utility functions of the form

$$h_0 = h_0(y, p_0, c_0; b), \quad (2a)$$

$$U_0 = U_0(y, p_0, c_0; b). \quad (2b)$$

Disposable income of migrants is  $y$  and  $p_0$  is the price of housing in serviced neighborhoods. Those residing in unserviced neighborhoods have, respectively, demand functions for housing and an indirect utility of the form

$$h = h(y, p, c; b) \quad (3a)$$

$$U = U(y, p, c; b). \quad (3b)$$

For the same housing price  $p$ ,  $U_0 > U$ , given  $c > c_0$ . To equilibrate utility across the two types of neighborhoods,  $p_0 > p$ , and from (2b) and (3b), we have

$$p_0 = p_0(p, y, b, c, c_0). \quad (4)$$

Equation (4) underlies the hedonic regressions reported earlier for Sao Paulo and Rio de Janeiro, examining within locality differences in relative rents based on type of service. Across communities, 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 equilibrium conditions: demand equals supply in the locality housing market and utility equalization for all migrants across localities within

the urban area.

### 3.1.1 Housing Demand Equals Supply

For the condition of 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 rising cost of bringing extra land into production of housing services. The number of serviced sites in the informal sector is a policy choice of existing residents, so the supply margin is the unserved sector with price  $p$ . Hence the supply specification  $H^s(A, p)$ . Summing the individual housing demands of the  $L_0$  serviced people (2a) and of the  $L - L_0$  unserved people (3a), using equation (4) for  $p_0$ , we have

$$L(b \cdot h(y, p_0, c_0; b) + (1 - b) \cdot h(y, p, c; b)) = H^s(A, p), \quad (5a)$$

$$b \equiv L_0/L. \quad (5b)$$

In (5b),  $b \equiv L_0/L$  is the proportion of migrants served, which is 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 rises as the total number of migrants  $\bar{L}$  to the urban area increases, with an inverse supply function of the form  $f(\bar{L})$ ,  $f' > 0$ . Equating utility of the marginal unserved migrant in our locality,  $U = U(y, p, c; b)$ , to this inverse supply, we can solve out for the locality housing price-level to get

$$p = p(\bar{L}; y, c, b). \quad (6)$$

Substituting for  $p$  from (6) into (5a), we get

$$L = L(b, A, \bar{L}; y, c_0, c). \quad (7)$$

Using urban area fixed effects (conditioning on  $\bar{L}$ ), we will estimate a version of (7), 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 because service levels are an externality. Thus, relative servicing is a policy instrument for localities to encourage or discourage in-migration. We also note that they affect overall housing demand and hence price level within the locality via capitalization and the proportions of serviced versus unserved housing ( $p_0$  and  $b$  versus  $p$  and  $(1 - b)$  on the left-hand side of (5a)).

### 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 (7) for each locality, noting that, for urban area  $i$ ,  $\bar{L}_i = \sum_{j,j \in i} L_j$ , where  $j$  indexes localities. Given the definition of  $\bar{L}_i$  and an equation (7) for each locality within the urban area, in principle we can solve for a system of equations for each locality where

$$L_{ji} = L_{ji}(\mathbf{b}_i, \mathbf{A}_i; y_i, \mathbf{c}_{0i}, \mathbf{c}_i), \quad j = 1, \dots, n_i, \quad (8)$$

$n_i$  is the number of localities in urban area  $i$ ,  $\mathbf{b}_i$  is the vector of service ratios of localities in urban area  $i$ ; and  $\mathbf{A}_i$  is the vector of land supply endowments for migrants in each locality in urban area  $i$ . Income levels and costs of publicly provided services are indexed by  $i$  to indicate they vary by urban area. Unit costs of services could vary by localities within an urban area, in which case we have vectors,  $\mathbf{c}_i$  and  $\mathbf{c}_{0i}$ . Finally, the function is also indexed since its form will vary with the number of localities,  $n_i$ .

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

### 3.3 The Strategic Choice of Servicing for Migrants

Existing local residents of locality 1 of a representative urban area choose  $b_1$  to maximize utility. We use a reduced-form specification of preferences,

$$V(y(\bar{L}), b_1, L_1; Z_1), \quad V_y, V_{b_1}, y_{\bar{L}} \geq 0, \quad V_{L_1} < 0. \quad (9)$$

For the first argument, we allow more migrants,  $\bar{L}$ , to the overall urban area to increase incomes of existing urban area residents,  $y(\bar{L})$ . This could reflect scale effects in urban

area total employment and/or the labor substitution effects in production where more low-skilled migrants raise the productivity of existing high-skilled residents. It is also meant to capture any urban area level diseconomies, such as generalized congestion on intra-urban area highways from having more workers in the urban area, although urban area scale effects in net here are modelled as positive. For migrants, we ignore these scale economies, although they are easy to add back in. The second term, where  $V_{b_1} > 0$ , reflects positive externalities from better servicing of migrants in the locality. However, the third term, where  $V_{L_1} < 0$ , implies that while residents of locality 1 want migrants to the urban area, they do not want them in their own locality. This could reflect local congestion considerations or simple prejudices against migrants.  $Z_1$  are other characteristics of the locality, such income of existing residents.

Existing residents choose  $b_1$  to maximize (9), 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

$$V_y y_L \left( \sum_{j,j \in i} \partial L_{ji}(\cdot) / \partial b_1 \right) + V_{b_1}(\cdot) + V_{L_1}(\cdot) \partial L_{1i} / \partial b_1 = 0, \quad (10)$$

where the  $\partial L_{ji} / \partial b_1$ 's are calculated from equation (8). Using this first order condition and equations (8), we can define

$$b_{1i} = b_{1i}(\mathbf{b}_{-1i}, \mathbf{A}_i, \mathbf{Z}_i; y_i, \mathbf{c}_{0i}, \mathbf{c}_i), \quad (11)$$

where  $\mathbf{b}_{-1i}$  is the vector of service levels in other localities in the urban area. Equation (11) will be the basis for estimating (1) strategic interactions and (2) how characteristics of existing localities influence policy choices. For strategic interactions, by differentiating equation (11), we can solve for  $db_{1i} / db_{ji}$ , or how locality 1 changes its service levels in response to a different offering in locality  $j$ . This gives a test of strategic interactions *per se*. But of greater interest empirically, as suggested in Figures 1b and 2, will be how different localities strategize according to whether they are higher-income versus lower-income or larger versus smaller.

### 3.4 A Simple Example

Consider an urban area with two localities, 1 and 2. Assume migrant house size is fixed and invariant to price. In any locality, serviced migrants have utility  $x + g_0^\gamma + b^\delta$ , where income to be spent on  $x$  is  $(y - p_0 - c_0 g_0)$ . Serviced migrants choose  $g_0$  to maximize utility, so that  $g_0 = (\gamma c_0^{-1})^{1/(1-\gamma)}$ . Thus, their quasi-indirect utility function is  $y - p_0 + f_0 + b^\delta$ , where

$f_0 \equiv (1 - \gamma)(\gamma c_0^{-1})^{\gamma/(1-\gamma)}$ . Correspondingly, unserved migrants have utility  $y - p + f + b^\delta$ . Equating utilities, we can solve  $p_0 - p = f_0 - f > 0$ , given  $c > c_0$ .

Housing supply to the community is given by  $Ap$ , where equating that to demand,  $L$ , we know  $p = LA^{-1}$ . Labor supply to the urban area is given by  $\bar{L}$ , so equating utility of the marginal unserved migrant,  $y - p + f + b^\delta$ , to  $\bar{L}$  and using  $p = LA^{-1}$  from local housing demand equals supply, for locality 1, we have  $L_1 = (y + f)A_1 + A_1b_1^\delta - A_1\bar{L}$ . Using the corresponding equation for community 2 and  $\bar{L} = L_1 + L_2$ , we can solve

$$\begin{aligned} L_1 &= a_1(y + f) + a_1(1 + A_2)b_1^\delta - a_1A_2b_2^\delta, \\ L_2 &= a_2(y + f) + a_2(1 + A_1)b_2^\delta - a_2A_1b_1^\delta, \\ a_s &\equiv A_s/(1 + A_1 + A_2), \quad s \equiv 1, 2. \end{aligned} \tag{8a}$$

Suppose an existing resident's utility is given by  $C + b_1^\delta - Z_1L_1^\phi$ ,  $\phi > 1$ , where the effect of immigration to the urban area on incomes is ignored. Maximizing with respect to  $b_1$ , where from (8a) we calculate  $\partial L_1/\partial b_1$  and  $\partial L_2/\partial b_1$ , we get with rearrangement

$$b_1^\delta = A_2/(1 + A_2)b_2^\delta + \beta_1 Z_1^{1/(1-\phi)} + C_1, \tag{11a}$$

where  $\beta_1$  and  $C_1$  are locality 1 specific constants (e.g.  $\beta_1 \equiv \phi^{1/(1-\phi)} a_1^{\phi/(1-\phi)} (1 + A_2)^{\phi/(1-\phi)}$ ). In equation (11a), strategic interactions are positive so there is a ‘‘race to the bottom’’, where if locality 1 lowers service levels, locality 2 follows suit.

It is easy however to construct examples where  $db_1/db_2 < 0$ . Suppose we rewrite existing residents' utility as  $(L_1 + L_2)^\varepsilon + b_1^\delta - Z_1L_1^\phi$ . Second, we make the localities symmetrical to keep the example simple. The first-order condition yields  $a\varepsilon(L_1 + L_2)^{\varepsilon-1} + \delta b^{\delta-1} - \phi a(1 + A)Z_1L_1^{\phi-1} = 0$ . To find  $db_1/db_2$ , we differentiate totally with respect to  $b_1$  and  $b_2$  to get

$$\begin{aligned} db_1/db_2 &= t_2/(-t_1), \\ t_2 &\equiv a^2\delta b_2^{\delta-1}[\varepsilon(\varepsilon - 1)\bar{L}^{\varepsilon-2} + Z_1\phi(\phi - 1)A(1 + A)L_1^{\phi-2}], \end{aligned} \tag{12}$$

where  $t_1$  is a collection of terms that must be negative from the second-order condition on the original choice of  $b_1$ , so the denominator of (12) is positive. Thus the sign of  $db_1/db_2$  is that of  $t_2$ . As in the prior example, if we assume a high distaste for migrants in the own community, so  $\phi > 1$ , then  $t_2$  and strategic interactions may be positive. However, assuming  $\varepsilon < 1$ , so there are not super scale and substitution effects, there are a variety of values of  $\phi$

where strategic interactions are negative. The intuition is straightforward. Rather than a race to the bottom, if one locality reduces servicing, another may raise servicing in order to attract more migrants to the whole urban area.

Equations (11) and (11a) will form the basis for tests of strategic interactions in the fashion of the public finance literature. Implementation is discussed below.

### 3.5 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. We may have both high-skilled migrants from outside the urban area and movements of existing high- and low-skilled residents across localities. While we will estimate overall locality household growth equations based on equation (7) (or  $L_1 = (y + f)A_1 + A_1b_1^\delta - A_1\bar{L}$  in the example in Section 3.4 above), we will also separately estimate growth equations for high- and low-skilled households. We introduce other considerations, such as differing tastes across localities among initial residents concerned with inequitable provision of public services as reflected in their voting preferences. Note the general specification of preferences of existing residents could incorporate the idea that migrants may be a fiscal burden or asset to existing residents of a locality.

## 4 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). There are two initial issues. First, as soon as we move beyond simple quasi-linear specifications of tastes and technology, the relationships governing  $b_1$  will be highly non-linear, with any locality's strategy interacted with all other localities' housing supply attributes and strategies. Second, across urban areas, there are different numbers of localities; this generates different forms to the  $b_1$  equation. Since there are many localities in large urban areas, we can have incredibly complex reaction functions as well as dampened strategic interactions as we move towards perfect competition. These issues plague the entire literature on estimation of local strategic interactions.

The ad hoc solution in formulating reaction functions is to decide that, for locality 1, (a) some localities are in more direct competition than others, and (b) rather than explore heterogeneous responses to other localities, other localities are represented by an index in a

simple linear specification. The standard formulation as reviewed in Brueckner (2000) is

$$b_{1i} = \beta \sum_{j \neq 1, j \in i} w_{1j} b_j + Z_1 \phi + \varepsilon_1, \quad (13)$$

where  $\sum_{j \neq 1, j \in i} w_{1j} b_j$  is a weighted sum of all other localities' choices in the urban area. Weights are typically chosen based on no explicit model. However, in our example in (11a), if we set  $\delta = 1$ , the model suggests we should weight by a land supply measure, where weights on competitors'  $b_j$ 's should increase as supply,  $A_j$ , does. We use the inverse of population density for the locality, as a weight indicating greater availability of land supply. In the literature (e.g., Case, Rosen, and Hines (1993) and Besley and Case (1995)), most weighting is based on spatial or economic proximity, in our case within an urban area, which is outside the specified model, although the intuition for proximity weighting is that a locality is in more direct competition with near neighbors. We utilize land supply weights since they arise in the model, but also discuss results using proximity weights and equal weights within the urban area.

We note that, while testing for strategic interactions is important, we are most interested in heterogeneity in the setting of  $b$ 's, or the role of locality conditions and the estimated  $\phi$  in (13). Writing (13) for all localities, we have

$$\mathbf{b} = \beta \mathbf{W}' \mathbf{b} + \mathbf{Z} \phi + \boldsymbol{\varepsilon}, \quad (13a)$$

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 an urban area (row) are normalized to sum to 1.

There are several issues in estimation of (13). First, by construction, since  $\varepsilon_{1i}$  influences  $b_{1i}$ , and since  $b_{1i}$  affects other localities' choices of  $b_{ji}$ ,  $\sum_{j \neq 1, j \in i} w_{1j} b_j$  is correlated with  $\varepsilon_1$ . Thus, OLS estimates are biased. In particular, if reaction functions are negatively [positively] sloped, a shock which causes one locality to raise its servicing will lead others to reduce [raise] their servicing. The absolute impact of reactions terms is therefore overstated. There are several possible solutions. The first is to use lagged values of covariates (Hayashi and Boadway (2001)), arguing that  $\varepsilon$  are uncorrelated overtime (especially in a context where we estimate with urban area fixed effects) and localities react in a lagged fashion to other localities' choices of  $b$ . We use this as one solution. A second solution is to instrument for  $\sum_{j \neq 1, j \in i} w_{1j} b_j$  with  $\sum_{j \neq 1, j \in i} w_{1j} Z_j$ , if we assume, for example, that time-lagged values of  $Z$ 's are exogenous (e.g., Fredriksson and Millimet (2002)). While in Section 5 we are

able to successfully instrument for individual  $b_{ji}$ , instruments for  $\sum_{j \neq i, j \in i} w_{1j} b_j$  have been unacceptably weak (with little impact on results), so we do not use that approach here.

A third solution for endogeneity of the  $b$ 's is to rewrite (13a) as

$$\mathbf{b} = (\mathbf{I} - \beta \mathbf{W})^{-1} \mathbf{Z} \phi + (\mathbf{I} - \beta \mathbf{W})^{-1} \boldsymbol{\varepsilon} \quad (14)$$

and estimate by maximum likelihood. The problem in estimating (14) is that it ignores spatial correlation of the error terms, which would arise, for example, from unobserved, correlated geographic factors across localities within an urban area that affect public infrastructure choices. Ignoring this fact in a specification like (14) would lead to biased estimates. Thus, 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

$$\mathbf{b} = (\mathbf{I} - \beta \mathbf{W})^{-1} \mathbf{Z} \phi + (\mathbf{I} - \beta \mathbf{W})^{-1} (\mathbf{I} - \Psi \mathbf{M})^{-1} \boldsymbol{\xi}. \quad (15)$$

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. Given urban area fixed effects are included in  $\mathbf{Z}$ , it is not clear what the sign of  $\Psi$  should be. For the  $\mathbf{Z}$ 's, we use lagged covariates to deal with issues of contemporaneous correlation between errors and covariates; however, policy choices,  $\mathbf{b}$ , are contemporaneous. In our case, maximum likelihood estimates of (15) are almost identical to the OLS estimates using lagged covariates (including lagged  $\sum_{j \neq i, j \in i} w_{1j} b_j$ ). Thus, with some degree of comfort, we will estimate versions of (13) that allow for more sophisticated strategic interactions than permitted by the linear model upon which (15) is based.

## 4.1 Results

We examine the determination of service levels for small houses (1–3 rooms) in which migrants are likely to live 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, both econometric and economic. We start with a base case where we are trying to determine a reasonable robust, econometric specification. Then we turn to different economic specifications, allowing for more sophisticated strategic interactions, and we have a detailed discussion of results. Finally we explore some counterfactuals.

### 4.1.1 Base Case

To look at econometric specifications, we start with a base case in Table 4a using a base set of locality characteristics (the  $Z$ 's), which are median household income in 1980, number of urban households in 1980, and the interaction between the two, with these variables all in logarithmic form. For strategic interactions, we use inverse density weights (for 1980) in summing opponent localities' strategic choices of service levels (the  $b$ 's) within the urban area, as suggested by the model.<sup>10</sup> We also report on other weighting schemes. Column 1 gives OLS estimates of equation (13), where own locality 1991 service levels react to opponents' 1980 service level choices. Column 4 repeats column 1 except it uses all 1991 covariates. Column 2 estimates the specification of equation (15) by maximum likelihood.<sup>11</sup> Column 2 estimates, as required, use own and opponents' 1991 service levels, given that (15) in principle solves the endogeneity problem of opponents' contemporaneous service choices. To mitigate other endogeneity issues, as in other columns, Column 2 uses 1980 covariates and inverse density weights. The inverse density weights apply to the  $\mathbf{W}$  matrix, although we experiment with alternatives. The  $\mathbf{M}$  matrix uses inverse distance weights between pairs of localities within the urban area as would be suggested by notions of spatial correlation. All specifications have urban area fixed effects, based on the model. Localities in the sample are at least 50% urbanized (so as to have an urban policy) in the base period.<sup>12</sup>

We start with an analysis of strategic interactions and estimates of the  $\beta$  coefficient. In the base linear model, all estimates of strategic interactions are negative. While a race to the bottom is an element of the model, the interpretation is that, in net, a rise in servicing in one locality leads other localities to reduce their servicing. The rise elsewhere helps bring more migrants to the urban area as desired for scale economy reasons, but lowering one's own servicing deflects these migrants to other localities in the urban area. As we will see below, we get more nuanced results when using non-linear specifications.

For the magnitude of strategic interactions, the OLS estimate in column 1 is -0.43. The corresponding OLS estimate in column 4, where we use contemporaneous measures for others' policy choices,  $\mathbf{W}'\mathbf{b}$ , shows the hypothesized direction of bias from using contemporaneous measures. The coefficient vastly overstates the absolute magnitude of strategic interactions. In column 2, the MLE estimate of  $\beta$  is similar to the OLS estimate, but is absolutely larger

---

<sup>10</sup>The choice of inverse density is meant to reflect greater weighting of opponent localities in the urban area that have more available land for development, and hence where new housing construction is likely to occur. It is the policies of these localities to which the own locality is most likely to respond.

<sup>11</sup>We perform the maximum likelihood estimation using James LeSage's econometric toolbox with the "sac" function in Matlab.

<sup>12</sup>There are also a handful of small localities where the actual number of small houses surveyed are fewer than 10; we exclude these to avoid noisy numbers on service levels.

at -0.53. We estimated many versions of the MLE model in column 2, in particular varying the weighting scheme for the  $\mathbf{W}$  matrix. Using equal weights within urban areas increased the absolute value of the  $\beta$  coefficient to -0.69 under OLS and -0.59 under MLE. Using inverse distance weights for the  $\mathbf{W}$  matrix yielded insignificant negative  $\beta$ 's, but there is the usual MLE problem of robustness and precision when using the same weighting scheme for both the  $\mathbf{W}$  and  $\mathbf{M}$  matrices. Given that both OLS and MLE estimates are sensitive to the choice of  $\mathbf{W}$  weights, we rely on the inverse density ones suggested by the model. These estimates suggest that an increase of 1 percentage point in the weighted average of small houses serviced in other localities of an urban area leads the own locality to service about 0.45 percentage points fewer small houses than it would otherwise, in order to deflect the increased number of migrants to the urban area to other localities.

Column 2 results have some particular features. First, the spatial correlation measure for error terms,  $\Psi$ , is negative, which may seem surprising, but recall we have urban area fixed effects. If urban area fixed effects are removed, the spatial correlation in errors becomes strongly positive. Second, estimates of  $\Psi$  are statistically weak here and more generally in variants of (15). Third, the standard errors of coefficient estimates in column 2 are less than in column 1, where column 1 allows for within urban area clustering, while column 2 specifies a common form to within urban area spatial correlation. Finally, in comparing column 1 and 2 results, as we will discuss later, the marginal effects of non-strategic interaction variables (the  $Z$ 's) are very close to each other. The results suggest to us that the OLS estimates in column 1 with lagged covariates are reasonable and close to relevant MLE estimates. Since we want to allow for more economically interesting strategic interactions than modelled in (15), we will tend to rely on OLS estimation with lagged covariates. This choice is enforced by the notion that MLE estimates are criticized in the literature as non-robust in a variety of circumstances (e.g., Conley (2008)).

Apart from strategic interactions, we are interested in how locality socioeconomics affects policy making. We believe that service provision in general is a normal good whose levels will rise with median locality income. 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 have a stronger strategic incentive to under-service small houses, as a means of deflecting migrants. They may have a stronger aversion to congestion or increased population density. Second, richer households may not want the children of low-income and low-educated migrants in local schools. Lastly, there may be fiscal reasons, such as the dilution of any property tax base, for richer localities to deflect low-income migrants. So while we might expect positive income and scale (in terms of

public service provision) effects, we expect the interaction between the two to be negative. All columns in Table 4a show positive income and scale effects, with a negative interaction term, as hypothesized.

What do the results in Table 4a suggest about locality preferences towards servicing of houses potentially occupied by low-income and low-educated migrants? Based on either column 1 or 2 in Table 4a, coefficients indicate that for localities at 1.5 standard deviations below mean income (at 8.9), a 3 standard deviation increase in size (4.2) increases servicing by about 0.14 or 14 percentage points (from a mean of 0.77), while at 1.5 standard deviations above mean income (10.1), the scale effect is close to 0 (0–0.03).<sup>13</sup> Similarly, at 1.5 standard deviations below mean size (at 7.2), a 3 standard deviation increase in income (1.2) increases servicing by just over 0.25 or 25 percentage points, while at 1.5 standard deviations above mean size (11.4), the scale effect is under 0.15. Table 4b explores this negative income-scale interaction further. Based on the Table 4a column 1 specification, we divide localities into separate size and income quintiles and then interact these, creating 24 cells relative to the base. As Table 4b shows, income and size effects both rise monotonically across quintiles. All interactive effects are statistically insignificant and not reported in the table except for the highest income quintile in the 4th and 5th size quintiles. These interactive effects are large, strongly diminishing the scale and income effects. So moving from the lowest income and size quintiles to the highest, ignoring interaction, raises servicing by 0.48 (48 percentage points); the interaction reduces that increase by 0.16.

Finally in Table 4a, column 3, we add covariates influencing locality choices to the column 1 specification. One 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 be a measure of preferences for more “leftist” representatives who might be more egalitarian (as revealed by actions in the subsequent democratic era). This indeed is associated with increased servicing. Higher own locality density also appears to increase servicing, which is intuitive since greater density would entail stronger negative externalities from poor water and sanitation conditions. These additional variables have little impact on the marginal effects of other covariates and we do not carry them through in other specifications that explore strategic choices in more detail.

---

<sup>13</sup>For example, the 0.14 number comes from  $(0.265 - 0.0261 \cdot 8.9) \cdot 4.2$

### 4.1.2 Locality Policy Setting

The evaluations and the linear model in the base case impose three key assumptions. First, strategic interactions are modeled as having the same form, regardless of how many localities there are in each urban area and how many competitors or opponents a locality faces. This is not what economic theory tells us. Second, interactions are linear, allowing, for example, for no interaction between what different competitors do. Third, the influence of the  $\mathbf{W}'\mathbf{b}$  variable does not vary by locality characteristics—high-income and low-income localities react to other localities’ strategic choices in the same fashion. We now explore an example in the data where we can relax these assumptions easily. We have 10 urban areas with 2 localities each and 13 with 3. Beyond that, we have a more limited sample of urban areas with specific numbers of localities: for example, there are 5 urban areas with 4 localities and there are a number of single urban areas with anywhere from 12 to 34 localities. So we experiment with the 2–3 locality urban area sample. We need a sufficient number of urban areas with any specific number of localities to distinguish effects by number of actors. Moreover, the interactive terms of strategic choices escalates with the number of actors.

We look at the sample of 23 urban areas that have either 2 or 3 localities. For the base specification, we have just  $\mathbf{W}'\mathbf{b}$  for 2 locality urban areas, since each locality interacts with just one other (and  $\mathbf{W}'\mathbf{b}$  is just the other locality’s choice). For 3 locality urban areas, we have  $\mathbf{W}'\mathbf{b}$  (which will now be a weighted average of the 2 opponents choices), but we allow the slope coefficient to differ since there are more players. In addition, for 3 locality urban areas, for each locality  $j$  we interact  $w_k b_k$  with  $w_l b_l$  to allow interactive effects. After estimating this specification of strategic interactions, we explore interacting opponents’ policy choice variables with own locality variables, in particular income, to see if high- versus low-income localities respond differently to opponents’ choices.

Results for the restricted sample are in Table 5. Column 1 of Table 5 repeats the specification of column 1 in Table 4a. Results are similar, although for this smaller sample, the point estimate for strategic interactions is larger in absolute magnitude and the coefficients on the  $Z$ ’s have sharper marginal effects. In column 2, we add in the new strategic choice variables. The one allowing 3 locality urban areas to have a different slope to  $\mathbf{W}'\mathbf{b}$  is negative but insignificant. More interestingly, the  $w_k b_k \cdot w_l b_l$  interaction term is strongly negative, indicating that in 3 locality urban areas, localities respond to the interaction of opponent’s choices. If both opponents raise their service levels, which helps attract more migrants to the urban area, that allows the own locality to deflect these migrants to the other localities by further lowering its own service levels.

Column 3 then interacts both the  $\mathbf{W}'\mathbf{b}$  and the  $w_k b_k \cdot w_l b_l$  terms with locality income.

There are three basic results. First, while the basic strategic interaction term  $\mathbf{W}'\mathbf{b}$  is negative throughout relevant income ranges (maximum income is below 10.6 in 1980), the negative strategic interactions decline sharply with income. That is, as other localities lower servicing, a richer locality raises its own service levels (to try to attract migrants to the urban area) by less than a poorer locality. Second, while the  $w_k b_k \cdot w_l b_l$  effect is negative over all income ranges (minimum income is above 8.2 in 1980), the negative reaction to enhanced servicing in both opponent localities is more muted as income rises. In general, richer localities are reacting less than poorer localities to others' choices. This means that in the net effect of reacting to others' choices—reflecting the trade-off between a race to the bottom (positive coefficient) versus attracting migrants to the urban area to enhance scale economy effects (negative coefficient)—the race to the bottom plays a more important role as income rises.

The final result is that, more generally, larger and richer localities display a strong tendency to under-service small houses. For example, in contrast to the results in Table 4, the column 3 estimates of Table 5 suggest marginal income and scale effects do not just decline with increases in income and scale, but become distinctly negative in larger and richer localities. To see this, we set  $\mathbf{W}'\mathbf{b}$  and  $w_k b_k \cdot w_l b_l$  at their averages for the relevant set of urban areas (0.303 for all localities for  $\mathbf{W}'\mathbf{b}$  and 0.236 for 39 localities in urban areas with 3 localities for  $w_k b_k \cdot w_l b_l$ ). Then marginal income effects are  $[(0.932 + 0.588 \cdot 0.303 - 0.569 \cdot 0.236) - 0.0982 \cdot \ln(\# \text{ urban households})]$ ; these income effects become negative by 0.5 standard deviations above mean size (10.0). Similarly, in this context, marginal scale effects are negative by one standard deviation above mean income. This suggests strong evidence of exclusionary behavior by localities that are both rich and large.

### 4.1.3 Robustness and Counterfactuals

Our localities are not always stand-alone municipalities, but are instead combinations of these political units. Over half of multi-municipality localities in our sample have over 85% of their urban population in the dominant municipality. We did rerun the basic models in Table 4 dropping the 25 localities where the dominant municipality had less than 65% of the urban locality population. There is almost no change in results, and we don't report them.

For counterfactuals, if we are correct about the nature of strategic interactions and incentives to deflect low-income migrants, then we should not see these effects in two situations. First, such modeling should not apply to servicing of large houses, which do not cater to low-income migrants, if we presume localities have no desire to deflect the rich. Higher-income households are generally well-served. For those without a central water connection, higher

income households can afford excellent private alternatives (deep wells, good filtering, and delivered water). The second situation is that, with democratization, the populist national government embarked on a wide-spread policy to upgrade slums and their servicing. Thus by 2000, we would expect to see much weaker effects. As Table 6 shows, there are no strategic nor income and scale interactions in servicing of large houses in 1991 and also none for small houses by 2000, based on the column 1, Table 4a formulation. Second, for the formulations (not reported) corresponding to those in column 2 Table 4a and in Table 5, income and size effects and their negative interactions are insignificant in both cases. For strategic interactions per se, there is no evidence of strategic interactions for the Table 5 formulation for 1991 large houses and limited evidence for 2000 small houses.<sup>14</sup> For the equation (15) model of column 2 Table 4a, while no income nor scale effects are present, there is evidence of strategic interactions for both counterfactuals. Nevertheless, we conclude that the OLS results are more compelling and that localities are not attempting to deflect high-income households in 1991 nor low-income households in 2000.

## 5 Effect of Service Provision on Locality Growth and Population Composition

We now turn to the effect of servicing decisions on locality population growth as presumably fueled by in-migration to the urban area. The specification is based on a linear version of equation (7), 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 then estimate is the within urban area allocation of migrants across localities. The basic estimating equation is

$$d \ln(L_{i,t}) = \beta A_{i,t-1} + \gamma b_{i,t-1} + \varepsilon_{i,t}. \quad (7a)$$

We look at local 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 conditions in the locality.

A key issue in estimation concerns the error structure. The urban growth literature (e.g., Glaeser, Scheinkman, and Shleifer (1995)) often takes the stance that (1) covariates are pre-determined and not affected by contemporaneous shocks that might induce growth, and (2) by looking at a growth equation, we have already differenced out time-invariant

---

<sup>14</sup>Only the  $\mathbf{W}'\mathbf{b}$  term in column 1 of Table 5 is significant for small houses in 2000. In other columns, all coefficients are completely insignificant.

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 and persisting sufficiently over time. Thus, 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 regulatory variable itself. Service supply today may be affected by past locality servicing, given bottlenecks in capacity expansion. Thus, low supply in 1991 may represent unmeasured good growth conditions from 1980–1991 for the locality, which caused a back-log in supply, and such growth conditions may persist into the 1990s. That is, high past growth is negatively correlated with current supply of water connections. Such influences will bias the estimated coefficient downward, understating the positive effects of good servicing on encouraging migration *per se*. The same issue relates to housing supply conditions—good unobservables driving locality growth in the past influence current housing supply conditions.

We focus on locality growth from 1991–2000. 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 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 expands 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 development of new agricultural crops for export, as well as the move from heavy industry based on state capitalism to lighter industry based on manufacture of consumer products (da Mata, Deichmann, Henderson, Lall, and Wang (2005)). This change in economic regimes will be part of an identification strategy. 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 1991–2000. We look for growth effects in the democratic era to see if poor servicing in 1991, which arose in the non-democratic 1980s, affects growth from 1991–2000.

## 5.1 Instruments

We need to instrument for two types of variables. First is the service variable and second are housing supply conditions in the locality.

### 5.1.1 Extent of Servicing

For water supply, we turn to geological variables, which affected historical private water supply in the form of wells. Geology influenced historical locality policies as to whether

to invest heavily in central water provision, even without exclusionary considerations. As water tables for wells have become strained, migrants at the margin without a central connection must obtain water by hauling from stand pipes or purchasing it privately (trucks, bags, bottles). What drove historical decisions about the extent of use of private wells? 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. Second, insolation affects the rate and variability of underground water replenishment, again a key issue in the viability of wells. 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 first-stage regressions, which show that these are strong instruments for servicing, with the expected effects.

We are using these geology variables to instrument for the externality variable: the extent of locality servicing as driven by historical conditions for getting good water service from private wells. An issue is whether these variables meet the exclusion restriction. The concern is that these variables could also be correlated with the unobserved cost of private alternatives today. We think the cost at the margin for private alternatives today is not associated with historical conditions for wells; rather, the cost at the margin is the cost of hauling water from stand pipes or purchasing it from water delivery trucks and stores. We discuss experiments with other instruments below.

### 5.1.2 Housing Supply Conditions

The other instruments are for variables relating to housing supply conditions, which, controlling for land, are the number of urban households, average education (influencing the demand for space), 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. Within urban areas, locality

economic bases have changed (an unobservable not captured in (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.

Use of historical instruments faces issues. There is a tension between going further back in time to break the persistence in relevant unobservables and weakening the strength of 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 give 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 draw from the following either using these variables 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;<sup>15</sup> (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 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. Combined with the geology variables, we experiment below with shorter and longer instrument lists.

---

<sup>15</sup>We experimented with replacing distance to Sao Paulo with latitude. Results are very similar, but specification tests favored the original set of instruments (both on strength and on orthogonality of instruments to error terms).

## 5.2 Effects of Servicing on Growth of Urban Households

Table 7 contains the basic results. Columns 1 and 2 contain OLS estimates, with shorter and longer lists of covariates. Columns 3 and 4 report 2SLS and LIML estimates for the preferred instrumental variables specification with a shorter covariate list. Column 5 shows that the variables omitted from the longer covariate list have zero coefficients in instrumental variables estimation. Column 6 repeats the column 4 LIML estimation with a shorter instrument list (designed for the shorter list of covariates); the corresponding 2SLS estimation for column 3 leaves coefficients unchanged.

### 5.2.1 Servicing

In OLS estimation, 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. 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 0.8, a one standard deviation (0.21) increase in servicing leads to an increase of 0.17 (for which the mean is 0.40) in the growth rate in the number of households 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. Water supply reduction in order to strategically retard locality growth below what it would have been in the absence of such choices 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). The partial  $F$  on the first stage regression for service levels is around 13. In a context where we are instrumenting for multiple interrelated variables including urban area fixed effects, this means instruments are not that strong. The LIML coefficients estimated to account for weaker instruments are generally larger, and in column 4 the coefficient is significant at just over the 5% level. However, LIML is more sensitive to the length of the instrument list, as column 6 reveals. The shorter instrument list brings the LIML coefficient more in line with the 2SLS estimates. For 2SLS, the 0.74 coefficient in column 3 is stable at 0.77 when estimated with the shorter instrument list in column 6.

An issue discussed above with our geology instruments for private supply conditions is whether they meet the exclusion restriction. Our problem is that with urban area fixed effects, we only have alternative instruments that are substantially weaker and perform poorly on under-identification tests. We experimented with a weaker instrument list for

servicing representing the costs of public provision (of no concern to the marginal unserviced migrant). Public supply often historically came from rivers. We instrument with whether the locality has a river, using that to replace the porous geology variable in the instrument list (including the interaction with insolation), and add in the altitude of the locality relative to the river (which captures the costs associated with pumping water upwards). The 2SLS coefficient on the extent of servicing is unchanged at 0.78. Of course, we must recognize that the existence of a river could also affect private supply costs. Removal of the porous geology and weather (insolation) variables as instruments leaves just locality historical political and economic conditions, which are weak instruments. The service variable coefficient does change from negative to positive, but the estimated coefficient is smaller and insignificant.

We perform two robustness checks for the growth equation. First, we reestimate the model in columns 3–6 of Table 7, dropping the 25 localities in which there are multiple municipalities where the dominant municipality has less than 65% of the locality urban population. Results are the same as in Table 7, so the inclusion of these localities has no impact on results. Second, as a service variable, we used the share of houses without land title that have a central water connection. We did not perform this same robustness check for strategic interactions because we lack 1980 data for this covariate. For the growth model, our instruments for the service variable are weak. The best case involves the short instrument list used in column 6 of Table 7, where the first-stage partial  $F$ -statistic for the service variable is 6.2. For this specification, the OLS coefficient of -0.16 becomes 1.82 under LIML but the standard error is 2.22.

### 5.2.2 Housing Supply Conditions

In Table 6, 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 instrumental variables estimation are what we expect. 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 an OLS coefficient in column 1 of -0.03 becomes -0.12 to -0.15 under instrumental variables estimation. Additional variables with OLS effects such as education (reducing supply) or share rural (increasing supply) have zero coefficients under instrumental variables estimation.

## 5.3 Composition Effects

So far we have looked at how under-servicing leads to a decline in overall in-migration. Because it is aimed at small houses, under-servicing 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 under-servicing 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 50% of the population in 2000, and those who have primary school or more. While we expect that poor servicing of small houses will affect growth of low-education households in a locality, there is the possibility it will also detract from growth of higher-education households. In the modeling in equation (9), we postulate that servicing of migrants is a positive externality for initial residents. Thus, we separately look at the effect of servicing small houses on growth of higher-education households.

The basic specifications are in Table 8. For housing supply, we control for base period overall density and household count of the relevant group whose growth we are investigating. Columns 1–3 deal with low-education household growth and columns 4–6 with higher-education household growth. Column 1 gives OLS estimates and columns 2 and 3 give 2SLS and LIML estimates. The deterrent effects of higher density and own group size are strengthened in moving from OLS to instrumental variables estimation as expected.

The focus is on the servicing of small houses. As expected, and as in the overall growth estimation, OLS coefficients for this variable are negative. In instrumental variables estimation, the service coefficient for the growth of low-education households is positive, very large, and now significant. A one standard deviation reduction in servicing (0.21) reduces the growth of low-education households by 0.17 for 2SLS, with even bigger effects for LIML estimation.

For high-education households between 1991 and 2000, under instrumental variables 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 somewhat noisier, being significant at just over the 5% level. Including a control for servicing of large houses occupied by high-education households (not shown), 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- relative to high-education households,<sup>16</sup> the coefficient on servicing is zero. Both that and the results in columns 2–3 and

---

<sup>16</sup>This ratio represents a form of differencing. We effectively difference what would be the separate equations for growth of low (L) and higher (H) education households:  $d \ln(N_{i,t}^k) = \beta^k \ln X_{i,t-1}^k + \gamma^k R_{i,t-1}^k + \varepsilon_{i,t}^k$ , where  $k = L, H$ . This has the advantage of differencing out location observables and unobservables whose effects are common to both groups. Where  $\gamma$  represents any differential in slope coefficient between low and higher education households, the estimating equation is then  $d \ln(N_{i,t}^L) - d \ln(N_{i,t}^H) = \beta_0 \ln(N_{i,t-1}^L/N_{i,t-1}^H) + \beta_1 \ln \text{density}_{i,t-1} + \gamma \ln(b_{i,t-1}^L) + \varepsilon_{i,t}$ .

5–6 suggest that poor servicing of small houses has adverse effects on growth of all education households, suggesting a strong negative externality for higher-education households.

## 6 Conclusions

While there is an extensive literature on the exclusionary policies of local jurisdictions (see Epple and Nechyba (2004) and Gyourko, Mayer, and Sinai (2006)), it has tended 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 into the informal housing market. Beyond enacting legal restrictions, localities can deny public infrastructure services such as water and sanitation 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 447 localities in 123 urban areas in Brazil between 1980 and 2000. These urban areas face exogenous employment shocks that attract migrants to the urban area, and migrants then make a decision of the locality in which to reside within the urban area. Localities themselves, however, may wish to limit population growth especially when migrants tend to be poor and low-educated. In the Brazilian context, demolishing informal housing is politically infeasible and there has been no recorded episode of such demolition since the early 1900s. Thus, localities make living conditions in the informal housing sector unpleasant as a means of deflecting migrants to other localities in the urban area.

We estimate the determinants of water provision to the types of houses in which the poor and low-educated primarily live (i.e., 1–3 room houses in 1991). We find evidence of strategic interactions—that localities adjust their service provision in response to corresponding policies in other localities of the urban area. We also 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, with the reductions in servicing dominating at the high-income end. Similarly, there is some evidence that richer localities are more likely to engage in a strategic race to the bottom.

We then estimate the effect of under-servicing small houses on locality population growth and composition. The findings suggest that under-servicing small houses leads to slower

growth in locality size. This slowed growth arises from both slower growth of low-educated households and slower growth of high-educated households. The implication is that externalities from under-servicing small houses are detrimental to attracting higher-educated households as well as to attracting lower-educated households to slum neighborhoods.

In sum, this paper provides evidence of intentional, strategic exclusion as an explanation for the existence of unserviced housing sectors (slums) in localities that have sufficient wealth and scale to provide basic infrastructure services to all houses in their jurisdiction. This intentional exclusion reduces locality population growth and alters its composition. Deflecting poor and low-educated households by making living conditions for them unpleasant also deflects the wealthy and high-educated households to localities where they do not confront the negative externalities of living near unserviced slums. Future research should focus on the effects of this intentional exclusion on social outcomes and economic growth of urban areas.

## References

- ADES, A. F., AND E. L. GLAESER (1995): “Trade and Circuses: Explaining Urban Grants,” *Quarterly Journal of Economics*, 10, 195–227.
- AU, C.-C., AND J. V. HENDERSON (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,” Mimeo, Paper prepared for Lincoln Institute of Land Policy.
- BESLEY, T., AND A. CASE (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,” Mimeo, Paper prepared for Lincoln Institute of Land Policy.
- BLACK, D., AND J. V. HENDERSON (1999): “A Theory of Urban Growth,” *Journal of Political Economy*, 107, 252–284.
- BRUECKNER, J. (2000): “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,” Mimeo, CASS.
- CASE, A., H. ROSEN, AND J. HINES (1993): “Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States,” *Journal of Public Economics*, 52, 285–307.
- CONLEY, T. (2008): “Spatial Econometrics,” in *The New Palgrave Dictionary of Economics*, ed. by S. Durlauf and L. Blume. Palgrave Macmillan, 2nd edn.
- DA MATA, D., U. DEICHMANN, J. V. HENDERSON, S. V. LALL, AND H. G. WANG (2005): “Examining Growth Patterns of Brazilian Cities,” World Bank Working Paper WPS3724.
- (2007): “Determinants of City Growth in Brazil,” *Journal of Urban Economics*, 62, 252–272.
- DAVIS, J., AND J. V. HENDERSON (2003): “Evidence on the Political Economy of the Urbanization Process,” *Journal of Urban Economics*, 53, 98–125.
- DOWALL, D. (2006): “Brazil’s Urban Land and Housing Markets: How Well are They Working,” Mimeo, Paper prepared for Lincoln Institute of Land Policy.
- EPPLE, D., AND T. NECHYBA (2004): “Fiscal Decentralization,” in *Handbook of Regional and Urban Economics*, ed. by J. V. Henderson, and J.-F. Thisse, vol. 4. North Holland.
- FREDRIKSSON, P., AND D. MILLIMET (2002): “Strategic Interaction and the Determinants of Environmental Policy Across US 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., J. SCHEINKMAN, AND A. SHLEIFER (1995): “Economic Growth in a Cross-Section of Cities,” *Journal of Monetary Economics*, 36, 117–143.
- GYOURKO, J., C. J. MAYER, AND T. SINAI (2006): “Superstar Cities,” NBER Working Paper No. 12355.
- HAYASHI, M., AND R. BOADWAY (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., AND A. KUNCORO (1996): “Industrial Centralization in Indonesia,” *World Bank Economic Review*, 10, 513–540.

JEFFERSON, G., AND I. SINGHE (1999): *Enterprise Reform in China: Ownership Transition and Performance*. Oxford University Press, New York.

WILDASIN, D. (1991): "Income Redistribution in a Common Labor Market," *American Economic Review*, 81, 757–774.

Table 1: Share of Urban Households with Servicing -- Average Servicing Levels in Localities  
[Weighted by Urban Population by Decade]

	Percent with central water connection	Percent with central water connection [Weighted]	Percent with full service: electricity, water and sewer	Percent with full service: electricity, water and sewer [Weighted]	Number of localities with water [Full Service]	Total Number of Localities
<b>Average for Localities at Least 50% Urbanized by Decade</b>						
1970	51	58	25	25	226 [149]	258
1980	70	80	30	44	365 [214]	376
1991	84	91	36	51	428 [250]	429
2000	89	92	49	62	435 [359]	435
<b>1991 Breakdown</b>						
Suburban Localities	82	86	32	39	303 [155]	304
Own house & land	86	91	36	49	428 [250]	429
Own house, not land	74	81	23	25	428 [250]	429
Rent	89	93	37	51	428 [250]	429
Ceded, other	77	85	30	43	428 [250]	429
Low Educ Houses	80	86	30	40	428 [250]	429
High Educ Houses	87	92	38	50	428 [250]	429
Migrants: bottom 20%	76	81	25	32	428 [250]	429
Non-migrants: bottom 20%	80	86	29	40	428 [250]	429
<b>Average of Localities at Least 50% Urbanized in 1970: Constant Sample</b>						
1970	51	58	25	25	226 [149]	258
1980	76	82	38	46	255 [173]	258
1991	89	92	48	55	258 [188]	258
2000	92	93	61	66	258 [239]	258
<b>Average for Localities with No Service (i.e. &lt;10% of houses served) in 1970: Constant Sample for Water and Full Service</b>						
1970	2	3	1	1	0 [0]	32 [109]
1980	57	68	9	30	30 [29]	32 [109]
1991	81	87	20	44	32 [50]	32 [109]
2000	88	92	37	56	32 [90]	32 [109]

Note: Includes all urban households living in localities that are over 50% urbanized. 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 HHs service is larger than the unweighted, this implies that the mean level of servicing is higher in larger localities. Note that for the localities with no service, the samples differ for water and full servicing: there are 32 localities at least 50% urbanized in 1970 with no water service and 109 with no full service.

Table 2: Servicing of Housing by Housing Quality

	Percent with Water			Percent with Full Service		
	1980	1991	2000	1980	1991	2000
Low Quality Housing 1-2 rms. 1980; 1-3 rms. 2000	61	81	83	17	32	42
High Quality Housing 6-7 rms. 1980; 7-9 rms. 2000	86	95	92	54	64	63

Note: Includes all urban households living in localities that are over 50% urbanized and have a minimum of 10% service levels. The restriction to a minimum of 10% service is meant to exclude those localities where there is in fact no service but where some households may have mistakenly declared in the census that they had service, or localities where service is incipient

Table 3: Use of Local Land Use Regulations

	% of urban localities having regulation, N=373. [for center cities of multi-locality urban areas]	Ratio of migrant (from 91-00) to resident h.h.'s, in regulated vs. unreg. localities	Ratio of homeowners without title to those with, in reg. vs. unreg. localities
Local min. lot size > 125 sq meters, 1999	64 [67]	1.09	1.16

Table 4a: Locality Decisions on Service Provision: Reaction Functions  
(1980 Covariates with Urban Area Fixed Effects)

	(1)	(2)	(3)	(4)
	Inv. density weight	Inv. density weight	Col. 1, added covariates	Col. 1 with 1991 covars instead of 1980 covars
	OLS	MLE	OLS	OLS
$\sum_{j=1, j \neq i} w_{ij} b_j$	-0.433** (0.151)	-0.531** (0.092)	-0.473** (0.163)	-1.43** (0.175)
Ln(median h.h. income)	0.426** (0.123)	0.383** (0.113)	0.315** (0.106)	0.222** (0.096)
Ln(# urban h.h.)	0.265** (0.133)	0.240** (0.110)	0.204** (0.106)	0.089 (0.083)
Ln(median h.h. income)* Ln(# urban h.h.)	-0.0261** (0.013)	-0.0230** (0.011)	-0.0217** (0.011)	-0.0099 (0.011)
Share, anti-military vote in nat. legislative elections			0.180** (0.089)	
Ln (density 1980)			0.0312** (0.014)	
Spatial error correlation		-0.195 (0.122)		
N [Urban Areas]	278 [51]	278 [51]	277 [51]	292 [54]
R <sup>2</sup> within	0.28		0.31	0.49

Table 4b: Effect of Income and Size on Service Provision  
(Quintile Specification)

	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Ln(#hh)	0.096*	0.111*	0.142*	0.191**
Ln (median income)	0.162**	0.208**	0.227**	0.267**
Statistically significant interactions				
4th size * 5th income	-0.174**			
5th size * 5th income	-0.160**			

Table 5: Policy Responses on Service Provision to Small Houses

	(1)	(2)	(3)
$\sum_{j=1, j \neq i} w_{1j} b_j$	-0.584**	-0.768**	-6.244**
	(0.163)	(0.166)	(2.231)
Ln(median h.h. income)	0.963**	0.921**	0.932**
	(0.351)	(0.384)	(0.305)
Ln(# urban h.h.)	0.827**	0.773**	0.904**
	(0.416)	(0.469)	(0.414)
Ln(median h.h. income)* Ln(# urban h.h.)	-0.0891**	-0.0840**	-0.0982**
	(0.044)	(0.049)	(0.043)
3 locality urban area * $\sum_{j=1, j \neq i} w_{1j} b_j$		-0.288	
		(0.358)	
3 locality urban area * ( $w_{1j} b_j$ * $w_{2k} b_k$ )		-1.178**	4.688**
		(0.536)	(2.209)
Ln(median h.h. income) * $\sum_{j=1, j \neq i} w_{1j} b_j$			0.588**
			(0.242)
Ln(median h.h. income) *			-0.569**
3 locality urban area * ( $w_{1j} b_j$ * $w_{1k} b_k$ )			(0.226)
N [Urban Areas]	59 [23]	59 [23]	59 [23]
R <sup>2</sup> within	0.55	0.58	0.64

Table 6. Counterfactuals  
Lagged X's; Inverse Distance Weights

	Service provision to large houses, 1991	Service provision to small houses, 2000
$\sum_{j=1, j \neq i} w_{1j} b_j$	-0.219	-0.21
	(0.182)	(0.229)
Ln(median h.h. income)	0.114	0.235
	(0.117)	(0.135)
Ln(# urban h.h.)	0.0318	0.108
	(0.099)	(0.098)
Ln(median h.h. income)*	-0.0015	-0.013
Ln(# urban h.h.)	(0.0100)	(0.0132)
N [Urban Areas]	278 [51]	292 [54]
R <sup>2</sup> within	0.25	0.16

Table 7. Growth of Urban Households:  
Share Small Houses Serviced with Water Connection

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	2SLS	LIML	2SLS	LIML
Share small houses with water, 1991	-0.183** (0.088)	-0.186** (0.078)	0.744* (0.429)	1.27* (0.657)	0.746* (0.431)	.859* (0.508)
Ln(no. urban h.h's, 1991)	-0.032** (0.008)	0.0061 (0.022)	-.0112** (0.031)	0 (0.050)	-0.115 (0.063)	-0.119 (0.038)
Ln (land area)	0.0053 (0.010)	-0.015 (0.014)	0.060** (0.020)	0.087** (0.034)	0.062** (0.035)	.065** (0.024)
Average education in locality, 1991		-.027** (0.014)			0.000 (0.049)	
Share s h.h's rural in locality, 1991		0.305 (0.196)			-0.041 (0.502)	
Urban Area Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N [55 Urban Areas]	351	351	351	329	329	329
Sargan stat. p-value			0.94	0.99	0.87	0.97
Min. 1st stage part. F			13.2	13.2	13.2	14.3
Anderson-LR p-value			0.12	0.12	0.07	0.03

Instruments are: locality adult illiteracy rate 1970, adult illiteracy rate in rest of urban area 1970, Share non-military vote 1982, manufacturing to service employ ratio in rest of urban area 1970, manufacturing to service employ ratio in rest of urban area 1970\* ln(distance to Sao Paulo), adult illiteracy rate in rest of urban area 1970\* ln(distance to Sao Paulo), ln (no. of urban h.h's in rest of urban area 1970), share h.h.'s that are rural in rest of urban area in 1970, share porous geology, mean insolation, standard deviation insolation, share porous geology\* mean insolation, share porous geology\* standard deviation insolation .

a. Short instrument list locality adult illiteracy rate 1970, share non-military vote 1982, manufacturing to service employ ratio in rest of urban area 1970, manufacturing to service employ ratio in rest of urban area 1970\* ln(distance to Sao Paulo), ln (no. of urban h.h's in rest of urban area 1970), share porous geology, share porous geology\* mean insolation, share porous geology\* standard deviation insolation.

Table 8. Growth in Different Education Households, 1991-2000  
 $[\ln(\# \text{ xx educ h.h.})(t) - \ln(\# \text{ xx educ h.h.})(t-1)]$

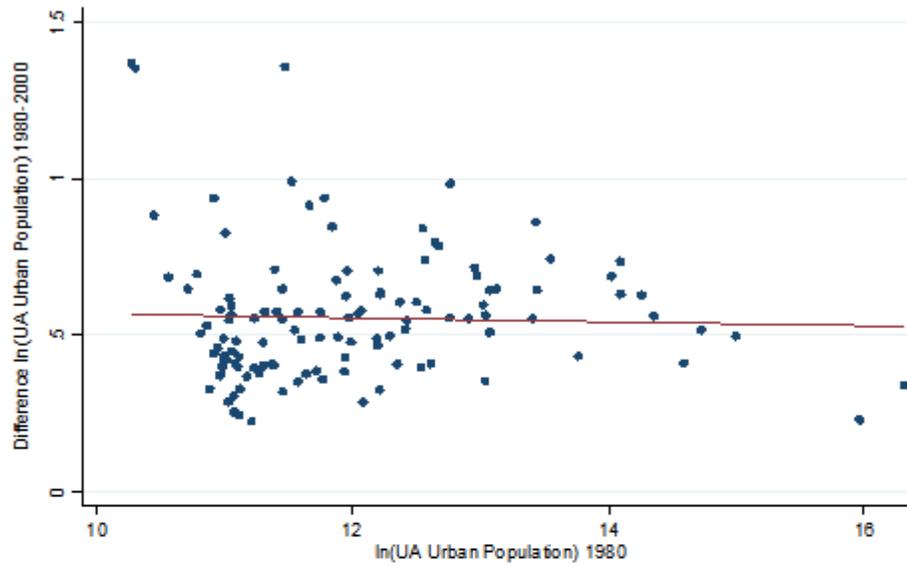
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	2SLS	LIML	OLS	2SLS	LIML
Share small houses with water, 1991	-0.170*	0.797**	1.59**	-0.056	1.05*	1.68*
	(0.093)	(0.296)	(0.644)	(0.097)	(0.577)	(0.873)
Ln(density, 1991)	0.014	-0.051**	-0.096**	-0.019*	-0.019	-0.116**
	(0.013)	(0.018)	(0.038)	(0.011)	(0.021)	(0.047)
Ln (#low educ. h.h.'s, 1991)	-0.035**	-0.061**	-0.085**			
	(0.011)	(0.018)	(0.027)			
Ln (#high educ. h.h.'s, 1991)				0.031	-0.082**	-0.036
				(0.010)	(0.027)	(0.029)
Urban Area Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N [Urban Areas]	351 [56]	329 [55]	329 [55]	351 [56]	329 [55]	329 [55]
Sargan stat. p-value		0.72	0.85		0.95	0.99
Min. 1st stage part. F		14.6	14.6		14.6	14.6
Anderson-LR p-value		0.12	0.12		0.12	0.12

Instruments are: locality adult illiteracy rate 1970, adult illiteracy rate in rest of urban area 1970, Share non-military vote 1982, manufacturing to service employ ratio in rest of urban area 1970, manufacturing to service employ ratio in rest of urban area 1970\* ln(distance to Sao Paulo), adult illiteracy rate in rest of urban area 1970\* ln(distance to Sao Paulo), ln(no. of urban h.h.'s in rest of urban area 1970), share h.h.'s that are rural in rest of urban area in 1970, share porous geology, mean insolation, standard deviation insolation, share porous geology\* mean insolation, share porous geology\* standard deviation insolation.

Figure 1. Urban Area and Locality Population Growth 1980-2000

a) Urban areas [slope coefficient (standard error) of -0.00644 (0.017);  $R^2=0$ ]

Urban Area, Urban Population Growth 1980-2000 and Initial Population in 1980



b) Localities [slope coefficient (standard error) of -0.0898 (0.011);  $R^2=0.14$ ]

Locality Urban Population Growth 1980-2000 and Initial Population in 1980

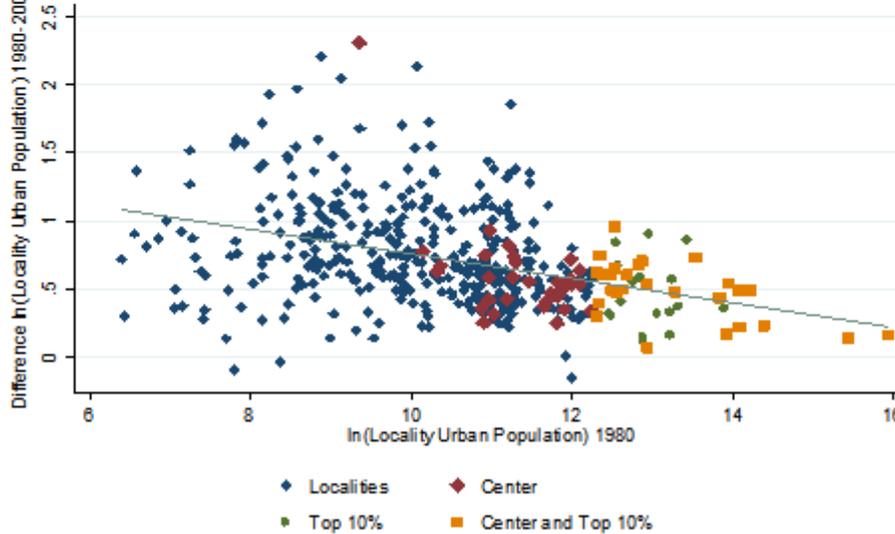
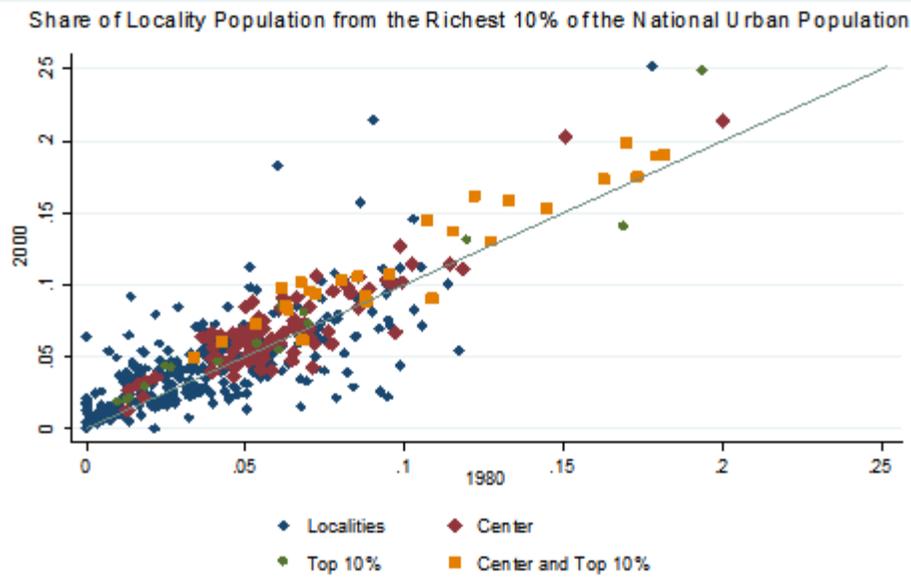


Figure 2. Share Rich in 1980 versus 2000



Regression (s.e. in parens.): Share Rich 2000=0.0113 (0.00208) + 0.943 (0.0365)\*Share Rich in 1980 + 0.112(0.0356)\*Share Rich in 1980\* dummy for center city.  $R^2 = 0.80$ . Sample is all localities with over 15,000 urban population and 50% urbanized in 1980; N=287. Excludes localities from the State of Bahia, for which there are no data on income in 1980.

## 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 run separately for each locality. We look at Sao Paulo and Rio de Janeiro municipalities in 1980. We do not have data on house prices, but we do have data on rental units, most of which are houses (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, 6 types of wall construction materials, 7 types of floors, 8 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 house attributes. To try 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 and then a reduced form where we use the typical summary measures—central water connection and full service (any electricity, central sewer, and connection to central piped water). In part, choices for the summary variables are driven by what data are available across census years. 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, but the results are suggestive.

Table A1. Rent Hedonics: Ln(Rent 1980)

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

## Appendix B

Table B1. Summary Statistics in 1991

	Mean	Std Dev	Min	Max
Share small house, water connection	0.74	0.21	0.021	1
Ln(# urban h.h. in locality)	9.5	1.47	5.8	14.7
Growth rate: # h.h.'s 91-00	0.402	0.175	-0.066	1.05
Growth rate: # low ed h.h.'s	-0.154	0.231	-1.04	0.404
Growth rate: # higher ed h.h.'s	0.697	0.318	-0.336	1.9
Ln (median h.h. income)	7.1	0.381	6	7.9
Ln (density)	5.44	1.63	0.85	9.4
Ln (#low educ hh)	8.49	1.35	4.9	13.3
Ln (#high education hh)	7.42	1.81	3.2	13.5
Avg schooling in city	4.74	1.11	2.4	8.8
Share h.h.'s rural rest of UA	0.26	0.188	0.03	0.91

Table B2. First Stage Regressions for Table 7. 1991 variables. N=330

	Share small houses, central water	Ln(# urban h.h.)	Average education	Share h.h. rural
Ln(land91)	-0.021* (0.011)	0.349** (0.059)	0.079** (0.030)	0.021** (0.008)
Illiteracy70	-0.0068** (0.001)	-0.074** (0.008)	-0.077** (0.007)	0.0045** (0.001)
Illiteracy rest UA 70	-0.027 (0.017)	0.0055 (0.130)	0.164 (0.119)	0.014 (0.012)
Illiteracy rest UA 70* ln(dist to SP)	0.0026 (0.002)	-0.014 (0.017)	-0.029* (0.016)	-0.0011 (0.002)
Manu-service employ ratio, rest UA	-0.123 (0.301)	2.63** (1.100)	-0.287 (1.340)	-0.036 (0.081)
Manu-service employ ratio, rest UA*ln(dist to SP)	0.012 (0.055)	-0.429** (0.206)	0.136 (0.241)	0.011 (0.015)
Ln(# urban h.h., rest UA) 70	-0.091** (0.038)	-0.942** (0.389)	-0.282** (0.136)	0.032 (0.029)
Share h.h. rural in rest UA 70	-0.03 (0.116)	1.48 (1.670)	2.32** (0.939)	-0.32 (0.148)
Share non-military vote, national. leg. elections 1982	0.165* (0.100)	2.03** (0.455)	0.045 (0.268)	-0.177** (0.061)
% geology porous	-0.489** (0.238)	-0.624 (1.440)	-0.712 (0.968)	-0.246 (0.210)
% geology porous*insolation mean	0.332** (0.108)	0.246 (0.713)	0.368 (0.536)	0.144 (0.117)
% geology porous*insolation standard deviation	-0.592** (0.191)	-1.31 (1.190)	-1.04 (1.020)	0.068 (0.175)
Insolation mean	-0.082 (0.206)	0.103 (1.010)	0.324 (0.735)	-0.214 (0.166)
Insolation standard deviation	0.164 (0.459)	6.69 (5.530)	4.11** (1.240)	-1.18** (0.538)
<i>F-statistic</i>	14.5	127	76	15.4