UN/POP/PD/EGM-URB/2008/03
31 October 2007

UNITED NATIONS EXPERT GROUP MEETING ON POPULATION DISTRIBUTION, URBANIZATION, INTERNAL MIGRATION AND DEVELOPMENT
Population Division
Department of Economic and Social Affairs
United Nations Secretariat
New York, 21-23 January 2008

THE EFFECTS OF EXCLUSIONARY HOUSING POLICIES ON URBAN DEVELOPMENT: EVIDENCE FROM BRAZIL*

*Leo Feler and Vernon Henderson***

*The views expressed in the paper do not imply the expression of any opinion on the part of the United Nations Secretariat.
** Brown University, USA.

Preliminary, incomplete draft: not for quotation

A. INTRODUCTION

In developing countries during periods of rapid urbanization, cities often house significant portions of their populations in informal housing sectors. From the 1950's through the 1990's, 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 African cities or the development of "urban villages," which house migrants in Beijing and other Chinese cities. Informal housing sectors may be characterized by varying degrees of insecurity of tenure, but perhaps most critically, they tend to be cut off from basic infrastructure services such as central water and sewerage. That makes living conditions unpleasant, unhealthy and expensive.

While unserviced housing developments may reflect failure in governance and the low incomes of residents, we ask whether, to some extent, the development of such districts may be strategic and intentional. Provision of bad living conditions for migrants is a way of existing residents resisting in-migration to a city, particularly of low income migrants. Such restrictions have immediate effects: inequality in living conditions, development of unhealthy neighborhoods and cities with high negative externalities, and restrictions on the supply of low-skilled labor to the city which may affect local economic growth. They also have implications for future policies as countries develop which are designed to make cities more "livable," or to engage in catch-up investment. Catch-up can be costly: building water and sewer mains long after development of dense neighborhoods can be very costly, requiring extensive spatial reconstruction of neighborhoods.

To better understand this process, we study Brazil over the last three decades. We examine (1) public infrastructure investment and land use regulation in localities in Brazil and (2) the impact of such policies on locality population growth and social composition. While we find evidence consist with strategic behavior, the impact of such policies on social composition of cities has a surprising aspect. In the end, under-servicing of informal housing sectors may also impact the location decisions of the rich, due perhaps to the negative externalities generated by unserviced slums.

Policies designed to spatially segment the population through inhibiting in-migration of, say, low-income people to particular communities is a well studied subject in the developed country literature. There are the fundamental insights of the Tiebout model (see review by Epple and Nechyba, 2004), analyzing spatial stratification of the population, whereby consumers stratify to live in communities with others with similar demands for public service levels; but this literature ignores the effect of stratification on production and income opportunities. More recently, Gyourko Mayer and Sinai (2004) analyze the development in the USA of what they call "superstar" cities. These are cities where population growth has slowed to a crawl; the share of the population from higher income groups is high and potentially increasing over time; and there is "excess demand to enter the city" as evidenced by rapidly rising housing prices relative to the nation. Such super-star cities may be endowed with high levels of natural amenities, often a natural harbor and attractive shoreline around which a city is built, attracting migrants. The presumption is that super-star cities have imposed strict land use regulations that inhibit further residential development which, if it occurred, would dissipate through congestion, the benefits natural amenities offer.

The situation in developing countries differs in two respects. First, an underlying presumption is that certain cities in developing countries have been favored, not necessarily with natural amenities, but as a policy initiative by national governments in terms of capital market allocations for industry, provision of public services for incumbent residents, licensing for export, FDI investment, imports, and government investment in state capitalism. The literature makes this point generally (Ades and Glaeser 199 , and Davis and Henderson, 2003) and then with examples drawn from Indonesia, and China, (Henderson and Kuncoro 1996, Jefferson and Singhe 200 and Au and Henderson 2006), as well as Brazil. This favoritism draws in migrants seeking job opportunities. If in-migration is unfettered, in the end, such favored cities would 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 may seek to restrict in-migration to the city, to halt this dissipation process. Attempts at restriction may be focused on exclusion of low-income migrants and apply less loosely or not at all to high-income migrants. Such an outcome might be a result of conflicting interests at city, state, and national government levels. While the city government may want to serve the interests of incumbent residents by limiting in-migration, national government officials may have a different political agenda, favoring certain cities and particular firms in those cities with subsidized investment and public services such as schooling, implicitly encouraging migration to these favored areas.

The second way the situation differs is that applying the notion of exclusion to developing countries is not straightforward. In many developing countries, exclusion from the formal housing sector, rather than halting city growth as in the USA, leads to development of an informal sector, which, unlike the USA, is "tolerated." By tolerated we mean it is not politically feasible to halt development of informal settlements and/or institutions are sufficiently weak that enforcement of a ban on informal settlements is not possible.

We develop these ideas looking at methods of exclusion and their impact on city populations and population composition in the context of Brazil over the last 30 years. Brazil has two types of informal sector housing markets. First are *favelas*, which were historically created, for example, by land invasions of government land or private land under title dispute. In principle such settlements are illegal, both because land use regulations are evaded and because the housing is on land not owned by those originally building, as well as subsequent owner-occupiers. Second are *loteamentos*, where developments do not meet zoning regulations, but are built on legally acquired land. However, after development, owners cannot obtain title because the housing does not meet zoning regulations. *Favelas* are an early phenomenon, often pictured in cities such as Rio as a response to in-migration pressure and lack of formal sector housing.

Loteamentos are a more recent development, supposedly in response to a national law in 1979 requiring 125 square meters of land as the minimum lot size for any construction. Since only 15 per cent of houses in Brazil are apartments, the law was aimed at single family homes. One view is that the law made formal sector housing unaffordable for lower- and many middle-income families. Since then individual cities have imposed even stricter minimum lot size requirements. As cities expanded after 1979, a substantial part of the increased housing demand was met by suburban developments that violated the national zoning law. Indeed, although suburban residents are 37.4 per cent of all urban households in major localities (see below), 43.2 per cent of homes without land title are in suburban areas in 2000.

A key aspect is that until the late 1980s and democratization, it was in principle "illegal" for cities to provide public infrastructure such as sewer and water in either type of settlement, providing an opportunistic excuse to deny or limit such provision. Thus, while cities cannot effectively halt informal sector development, they can contain it. A strategy in containment seems to be to deny basic public infrastructure services to such communities, with a twofold result—poor living conditions and the need to substitute in expensive private alternatives to public provision of basic services. We analyze under what conditions cities "under-provide" public infrastructure to poorer neighborhoods. Then we estimate the impact of such under-provision on city populations and population composition.

Section 1 of the paper discusses data and trends towards exclusion and the development of "superstar" cities in Brazil. Section 2 discusses a conceptual framework, to inform econometric specifications. Section 3 looks at the determinants of provision of public services. Section 4 estimates the impact of exclusion on locality population growth and population composition. Section 5 concludes.

1. Urbanization and public infrastructure in Brazil

This section provides an overview of patterns of growth and population composition of localities in Brazil, as well as information on the extent of servicing of communities with basic infrastructure. The section gives background information relevant to the development of hypotheses and estimations in the sections to follow.

The paper focuses on the post-1980 time period, in part because of data availability. 1980 is near the end Brazil's period of rapid industrialization and urbanization. Industrial development, which had focused on São Paulo and Rio in the post-World War II period starts to decentralize in the late 1970's, with substantial and ongoing industrialization of hinterland cities. This decentralization is facilitated by inter-city investments in transport and telecommunications, as well as agriculture developments in the north. As we will see, post-1980 urban areas everywhere have about the same growth rate; and, if anything, growth is more rapid in the north compared to the traditional region of growth—the southeast (de Mata et al., 2003). This change in urbanization and industrialization patterns will provide a basis for one instrumental variables strategy, as detailed later.

1.2 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 socio-economic information covering education, income, family structure, and migration. We also have information from on basic geographic and fiscal indicators. For land use regulation we have retrospective information. A census of local governments conducted in 1999 and in 2005 [IBGE, Perfil dos Municipios Brasileiros, 1999 and 2005] indicates whether cities had passed a minimum lot size zoning law in excess of the national standard of 125 sq meters by 1999. We know if cities have "zoning laws," meaning land use regulation under potentially a master plan allocating land to different uses and intensities. We also know if cities have a "parcel law," which would include minimum lot size restrictions. We know from retrospective questions posed of city administrators in 1999 and 2005 the date of the first such law passed in the city. Evidence (Biderman, 2007) suggests this information is flawed; in retrospective dating, cities tend to ignore/forget the earliest laws.

The effective level of local government in Brazil is the *municipio*, a unit akin to a USA county, with larger urban areas consisting of multiple *municipios*. Our unit of analysis in principle is the *municipio*, as well as 123 urban areas (defined as of 1991) consisting of either a single *municipio* or a collection of *municipios* (59 of the cases). However, the overtime analysis is complicated by the fact that some initial *municipios* split into more *muncipios*, and there were some recombinations as well. In particular, fiscal formulas after 1988 favored creation of new *municipios* (formulas included a lump sum component independent of size). To facilitate overtime analysis, we combine split *municipios* into "common denominator" ones, which we call localities (informally) or MCAs more formally (see da Mata et al., 2003). The 123 urban areas are composed of 447 localities [MCAs], which, in 1991, consisted of 659 *municipios* and, in 2000, 833 *municipios*. Urban areas are either defined as agglomerations by the Brazilians, or are single localities with over 75,000 people which are over 75 per cent urbanized in 1991. Since what is urban changes dramatically in the 447 localities from 1970 to 2000, especially the suburban localities, we often look at sub-samples of localities, imposing criteria such as requiring the locality to be at least 50 per cent urbanized in a given Census.

1.2 Patterns: urban growth, stratification, and superstar localities

Urban areas in Brazil experienced "parallel growth" from 1980-2000, meaning that small and large areas grew about the same rate, as a number of theories with or without a stochastic component predict (e.g., Black and Henderson, 1999 and Gabaix, 1999), although the dispersion of growth rates is larger for smaller urban areas, as pictured in Figure 1a. There is no evidence of "convergence," or relative mean reversion for urban areas. In contrast, in Figure 1b, localities experience significant mean reversion: bigger localities, often central city ones, grew at a slower rate than smaller localities. Overall, these differing patterns between urban areas and localities represent decentralisation within urban areas, where, with rising incomes and populations, urban areas spread out and suburbs develop, fueled by declining commuting costs with transport improvements which make central city locations less valuable.

But Brazil's pattern of decentralisation has one key feature, which differs from USA patterns. As in much of the world, the rich live predominately in the center cities and the poor in the suburbs. Why the difference? One reason may be that, unlike the USA, in most countries, funding for public education occurs at the state or national level. The rich by suburbanizing can't form exclusionary "clubs" offering independently funded, high-quality schooling; and thus, they may prefer the centre city with its lower commuting times to service intensive central business districts.

Table 1 gives data for 1970 and 1980 versus 2000 for centre cities and suburbs of the 59 multi-locality urban areas, with Rio and São Paulo centre cities and suburbs separated out. The table shows for each geographic construct what share of local households are in the richest 10 per cent of households (in the national urban sample) and what share are in the poorest 10 per cent. São Paulo and Rio centre localities have a very high ratio of rich to poor, formed by both high shares of rich and low shares of poor. While that ratio has declined with time as these centre localities have absorbed more poor, especially after democratisation, it is still very high. The suburbs of Rio and São Paulo show two distinct patterns. First, they also start with a high ratio of rich to poor which declines sharply overtime to almost 1. Second, suburbs have many small poor localities even though overall, their share of rich to poor is high. For example, while the weighted (by number of households) average of the per cent rich in 1980 is 12.1, the unweighted average across suburban localities is only 5.5.

For other cities, there is a distinct change over time. Centre cities start with a low ratio of rich to poor, although much higher than their suburbs. Since resistance to population growth occurs at the level of the *municipio*, older, often central city *municipios* are the most likely to suffer from over-crowding and most likely to desire to limit population growth, especially of low-income residents. So a strategy of central cities may be to deflect immigrants and lower income residents to their suburbs as well as to other cities.

In the data, centre cities have gained increasing relative shares of rich to poor. For other suburbs in other urban areas (than Rio and São Paulo), while there is some loss of share poor, they remain with a very low ratio of rich to poor in 2000.

More generally in urban Brazil, it seems that large localities, many being central city ones, continue to grow richer. In Figure 2 we plot the per cent rich in 2000 against 1980. We see that centre cities and other large localities have improved their share rich in general: the data points lie above the 45 degree line and the slope coefficient for central cities is 1.06 versus .94 for all localities. This means the increases for large localities of shares of rich in Table 1 are across the board.

Finally, pulling together information in prior figures, Figure 3 suggests elements of a superstar story. In the graph, the vertical line defines the (unweighted) mean locality population growth for 1980-2000 of localities and the horizontal line marks the mean (weighted) per cent rich, defined as the richest 10 per cent from the urban sample of localities. The upper left quadrant then are slow growing areas dominated by richer populations – these are mostly large localities, in the top 10 per cent by size. We interpret these as being favored localities that initially may have drawn in significant shares of poor, which have grown relatively slowly, and which have successfully increased their share of rich.

1.3 The informal sector and public infrastructure

An issue in the literature on Brazil (e.g., Dowall, 2007) concerns how to identify who lives in the informal sector. In the Census, there is a question filled out by census takers on whether people live in "irregular settlements," which is based on the "irregularity" of the surrounding configuration of housing, not on land ownership. Less than 5 per cent of households are irregular; researchers tend to dismiss the Census numbers as not capturing informality. For tenure mode in the 1991 and 2000 Census, for homeowners, there is also a question on whether they have title to their land. Although many households without true title may answer yes to having title, the question serves to identify a set of households and their circumstances whom we infer truthfully answered no. In 1991 about 8.2 per cent of urban households in significant size urban settlements (over 15,000 urban people- see below) live in owner-occupied housing for which they do not report land title, which corresponds to 13.5 per cent of owner-occupiers. As noted above, these numbers are presumed to be a lower bound on the magnitudes of those without formal sector land tenure. Brazil has undertaken a number of tenure regularisation programmes. What seems to be apparent is that even without land tenure, especially in *loteamentos*, people do not feel insecure about their holdings. Moreover, in regularisation programmes, reportedly, a number of participants fail to take the last step and register their land tenure once they would be able to do so.

A different approach is to define informality on the basis of lack of public infrastructure provision (Dowal, 2007). From the literature, it seems a key element is a central water connection, where in 1991, about 14 per cent of urban households 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 per cent of households do not have full service. To check on the validity of the presumption that such services are highly valued, we examined willingness pay for such services, via hedonic rent regressions for renters in São Paulo and Rio de Janeiro *municipios* for 1980, when relevant data are available. We find consumers are willing-to-pay about 20 per cent more for a rental unit with central water and an additional 20 per cent in São Paulo and 30 per cent in Rio for a unit with central sewer and electricity, in addition to water. Breaking out some of the components in the Appendix, it seems there is a very high premium on electricity (although even in 1980 it is virtually universally available); but central sewer still commands 9 and 18 per cent premiums in São Paulo and Rio respectively, over no sewage (with septic systems in these congested cities generating little *premia*).

In this paper, we do not attempt a definition of informality per se, partly because it seems very difficult to do so. We simply focus on the notion of exclusion through lack of servicing, whether housing is in the formal or informal sectors (in terms of title and meeting of land use regulations). Note the impact of, say, minimum lot size regulations on overall locality size and income composition is unclear, given it may facilitate the growth of the informal sector. Thus, the aspect of exclusionary tactics which we focus on is to not service housing with basic water and sewer, forcing residents to live in poor conditions and/or attempt to substitute in high priced private alternatives. Having said this, infrastructure provision or lack thereof is an issue with many dimensions. It is difficult to separate out deliberate withdrawal of services from: (1) geography which influences the need for public provision of sewer and water, as opposed to relying on wells and septic systems; (2) issues to do with city income and overall public sector demand for services at different stages of development; and (3) the political economy issue, that even without exclusionary motives, there is a local political-social mechanism determining what areas of a city first get upgraded services. How does servicing vary over time and people in Brazilian cities?

a. Provision of infrastructure services

Table 2a and 2b explore dimensions of servicing. To limit the problem of geography and initial up-front investment costs we look just at localities that are over 50 per cent urban and that provide services to at least 10 per cent of households. First, Table 2a shows how services in such areas have changed over the decades and marks the rapid improvement of services in urban Brazil. In 1970, only 53 per cent of households were connected to a central water system; but 2000 this was 89 per cent. For full servicing, where the shortfall relative to a water connection is lack of central sewer, in 1970, 22 per cent of households had full service while by 2000 this number is 59 per cent. For the year 1991, the rest of Table 2a explores spatial, housing tenure, and income differences in provision that persist in relative terms in all years. Suburban areas with their low population densities and lower incomes have poorer servicing than centre cities 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. The worst served people are those who report that they don’t own the land under their house. By definition these are informal housing sector residents, for whom, as noted earlier, until the late 1980s it was “illegal” to provide public infrastructure. Similarly poorly served are people living in “ceded” housing and “other,” such as employer provided housing. This category is dominated by ceded—people who are living rent free (but, in principle, not in employer-provided housing).

Finally in Table 2a, since we think of exclusionary mechanisms as being aimed at migrants, we compare the servicing of migrants with non-migrants. Migrants are households where the household head moved to the locality within the last 10 years. In the last two rows, we compare services for migrants with non-migrants for those in each group who are in the bottom 20 per cent of the national urban income. In terms of full servicing there is a distinct difference for full servicing in 1991 between migrants and non-migrants and for water as well in earlier years. But perhaps the main feature is that those in the bottom 20 per cent of households are very poorly serviced—migrant or not. While localities care about the rate of in-migration, today's migrants are tomorrow's non-migrants. Localities may be more 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. However localities can't discriminate on the basis of income per se, nor by migrants versus not migrants.

What localities can do is to not service the houses that low income people are likely to occupy, while servicing the houses that higher income people are likely to live in. In Table 2b, we examine the service levels for the smallest houses likely to be occupied by the poor: 1-2 rooms in 1980 covering the bottom 14.3 per cent of the house size distribution and one to three rooms in 2000 covering 16.5 per cent of houses. We compare these with houses for upper middle income folks: those with six to seven rooms in 1980 covering 21.5 per cent of households (below the top 11 per cent by size), and seven to nine rooms in 2000 covering 19.7 per cent of households (below the top 6 per cent). The focus is on the intra-period differences between high- and low-quality houses in terms of servicing. We look at the differentials overall and for central cities in particular, where as already noted central cities with their longer history of development and high densities are better serviced in general.

In Table 3b, in 1980, only 61 per cent of small houses in localities had water, while 86 per cent of large houses had water. More stark is the relative difference between full servicing: 17 per cent versus 54 per cent. Gaps for central cities are similar. By 1991, the water gap has diminished with economic development and perhaps due to democratisation. However the relative full service gap remains very large. In the section 3 and 4, we will use the servicing of smaller houses as our basic "exclusionary measure."

Land Use Regulations

In Table 3, we examine local land use regulations, over and above the national 1979 minimum lot size law. Most of these have been passed since 1988; and our data are from 1999 and 2005 as noted earlier. In Table 3, we list what fraction of significant size localities (over 15000 urban) by 1999 had passed a minimum lot size law in excess of 125 sq meters—that is, a minimum lot size law that is in excess of the national standard. In column 2, we show the fraction of urban households in 2000 that are migrant households in locally "regulated" areas (those who have more stringent minimum lot size law than national law). In column 3, we list the ratio of homeowners without title to those with in regulated versus unregulated localities. Note more migrants relative to non-migrants and more household without title relative to those with title tend to live in areas with minimum lot sizes in excess of the national standard. This might hint that localities which impose stronger zoning regulations are those subject to migration pressure. In other rows in Table 3, we also note corresponding numbers based upon retrospective information gathered in 1999 and 2005 as to what cities had "zoning laws," governing the usage of land and what cities had parceling laws (e.g., their own minimum lot size lot land use) by 1991 and then by 2000. There is no information on the restrictiveness of these codes. The fact that they apply more in communities with relatively high numbers of home owners reporting no title could reflect either exclusionary sentiments or some attempt to regularize the land market ex post.

2. Conceptualizing exclusionary behavior

In this section, we present a simple stylised model to analyse both urban growth and exclusionary behavior. The model will inform the specification of estimating equations. Consider a locality with an initial resident population which is facing an influx of immigrants. The supply curve of migrants to the city is $V^S(N_m),\ V_1^S \geq 0$, where V^S and N_m are respectively a migrant's alternative utility level and the total number of entrants, or migrants to the city. The supply curve could be upward sloping, as migrants who move from further distances or better opportunities in alternative locations demand higher returns to move to the city.

On the demand side, representing what welfare levels the city offers migrants, we assume critically that the marginal migrant is in the informal sector. The cost of housing services (net of land costs), including public utilities such as water and sewer is much higher in the informal (I) sector, compared to the formal (F). We denote that unit quality-constant cost as c, where $c_I >> c_F$. Included in the high cost in the informal sector are the high cost of private provision of water and sewerage, compared to public provision which enjoys high economies of scale. The indirect utility function of the marginal migrant to the informal sector is

$$V^I(w_m(N_0 + N_m; A),\ c_I,\ N_m^{NS}/N_m); \quad V_1^I > 0,\ \ V_2^I < 0,\ \ V_3^I < 0. \qquad (1)$$

$w_m(\cdot)$ is the real wage of migrants, which incorporates both nominal wages paid by employers which are increasing in city scale and the costs of commuting, which detracts from time available to work, following Duranton and Puga (2004) or Au and Henderson (2006). It can also include land rental costs, which rise along with commuting time as city size expands (Henderson and Venables, 2007).

In equation (1), real wages are presumed to initially rise as city employment expands with exploitation of scale externalities, but eventually at the margin these scale benefits are offset by higher costs of commuting and real wages start to decline with further increases in city population and employment. Thus real wages are an inverted-U shaped function of city employment; the usual presumption is an in a well behaved equilibrium in national labor markets is that cities operate on the downward sloping portion of that real wage curve (see below). In the wage function $w_m(\cdot)$, A represents key items such as local technological levels or knowledge accumulation (Moretti, 2004 and Black and Henderson, 1999) and geography such as the amount of land in the locality which affects land costs.

In equation (1), as already noted, c_I is the unit cost of constant quality housing services including utilities and welfare is declining in that unit cost. We could define c_I to be a function with geographical elements reflecting locations of water tables and rainfall. Finally, we argue that the welfare of the marginal migrant is declining in the fraction of migrants not served by public utilities, where N_m^{NS} is the count of migrants who are not served with public utilities. The idea that welfare is declining in the count not served is based on externalities; a larger number unserved results in more pollution and unhealthy living conditions, with poorer private water quality, higher rates of septic system failure and general contamination of soil and ground water, and greater spread of contaminants.

Equating utility demanded by migrants to that offered to the marginal migrant by the city, we have the implicit function $V^S(N_m) - V^I(w_m(N_0 + N_m; A), c_I, N_m^{NS}) = 0$, which can be solved to yield

$$N_m = N\left(N_0, A, c_I, N_m^{NS}\right). \qquad (2)$$

By simple application of the implicit function theorem rule, in (2) N_m is declining in all arguments except *A*, given prior assumptions (including the notion that we are operating on the downward sloping portion of the inverted-U real wage function). In writing (2) we are treating N_m^{NS} as an implicit policy variable to be determined by the locality's existing N_0 residents. As we will see momentarily, the community sets the number of migrants to be served, N_m^S; and, based on the analysis below, we presume that as N_m^S rises, N_m rises and N_m^{NS} declines.

Figure 4 illustrates the equilibrium level of migration, under two different circumstances. The initial level of residents is N_0 and the supply of migrants to the city starting from N_0 is depicted by the $V^S(N_m)$ curve. The $U(\cdot)$ and $V^F(\cdot)$ curves show the welfare of respectively initial residents and of migrants to the formal sector (or at least migrants receiving city services), as a function of city scale, without factoring in how differences in N_m^{NS}/N_m might impact this welfare (see later). $V^F(\cdot)$ is shown as lying below $U(\cdot)$, on the basis that migrants may have lower skills, or may reside in an informal sector even if they are fully serviced. Note also that migrants who are served will pay a rent premium relative to those not served, so their utility levels are equalized.

For migrants in the unserviced informal sector two scenarios are shown. In the first all migrants are unserviced and in the informal sector, so $N_m^S = 0$ and $N_m^{NS}/N_m = 1$. Their utility curve is given by $V^I(\cdot, N_m^{NS}; N_m^S = 0)$, shifted down relative to $V^F(\cdot)$ because of the high unit cost c_I and sharply downward sloping because of increasing numbers of unserviced residents and the externalities that they impose. City population is N^A, with $N^A - N_0$ migrants. In the second regime, the city services $\bar{N}_m^S$ migrants and the $V^I(\cdot, N_m^{NS}; N_m^S = \bar{N}_m^S)$ curve plots the welfare of the marginal unserviced migrant. Total city population is N^B, with $N^B - \bar{N}_m^{NS} - N_0$ unserviced migrants. Note as the number of serviced migrants rises, N_m^{NS}/N_m and N_m^{NS} both decline, but total migrants rise.

Initial city residents pick a N_m^S and implied N_m^{NS} to maximize their welfare, given by $U(w(N_0 + N_m), c_F, N_m^{NS})$ where $U(\cdot)$ is declining in N_m^{NS} for externality reasons. Optimising $U(\cdot)$ with respect to N_m^{NS} and assuming an interior solution, we get $(\partial U(\cdot)/\partial w)(\partial w/\partial N)(\partial N_m/\partial N_m^{NS}) + (\partial U(\cdot)/\partial N_m^{NS}) = 0$. Note the first term is positive, assuming $\partial w/\partial N < 0$ in equilibrium and $\partial N_m/\partial N_m^{NS} < 0$, while the second is negative. The second represents negative externalities for incumbents of under-servicing migrants, which we will argue in the empirical section are very strong.

We could think of added aspects to the welfare of initial residents such as political beliefs about inequitable provision of public services. We could also have the locality start as vastly under-populated operating on the rising part of the $U(\cdot)$ and $V(\cdot)$ curves, providing further incentives to service migrants so as to attract them to the city. And we could introduce tax considerations, where migrants are a fiscal burden or asset (pay taxes but get no services).

3. Determinants of locality infrastructure servicing and land use regulation

We have two types of restrictions: explicit land use regulations and implied restrictions on servicing of types of housing in which migrants might typically live. For land use restrictions, apart from the 1979 national law, our impression is that these only really start to play a role at the local level after 1988. The key restriction in the 1980s was the decision to service or not service areas of the city, particularly those "illegal" settlements, which violated the national law. We start by looking at provision of water and then the full range of services—water, sewer and electricity (where the last is almost universal) to houses likely to be demanded by migrants, in particular small (1-3 room) houses in 1991. This measure indicates the quality of services available to lower income migrants.

We hypothesize that the extent of servicing of small houses in time *t*, *S(t)*, is a reduced form function of demand and supply conditions, as well as strategic considerations. Demand conditions include normal income and externality considerations. On the supply side are a variety of cost considerations, including geographic conditions, institutional factors, economies of scale, and capacity constraint considerations. What we look for, at the moment, is evidence of outcomes that reflect strategic considerations, rather than just normal income and scale effects, in the provision of services to small houses intended for migrants, but not other houses. The formulation is in (3).

$$S(t) = S\left(y(t-1)),\ \ln(\mathrm{Z}(t-1)),\ \ln(\mathrm{Z}(t-1))^{*}\ln(\mathrm{y}(t-1));\ X(t-1) \right) \quad (3)$$

In (3), the first set of variables represents demand, scale, and strategic behavior. Normally, we would expect provision to rise with community income, as a regular income effect and to rise with community size due to scale economies. In the first three variables, *y(t-1)* is median income of the locality and *Z(t-1)* are attributes which might trigger strategic behavior, at higher income levels. One attribute is locality size; another is the fraction of households that are poor in the rest of the urban area and who are a threat to spillover into the own locality. While larger localities can exploit economies of scale in providing public infrastructure which gives a positive scale effect, as income rises, the marginal scale effect may become negative due to strategic exclusionary behavior, consistent with the "superstar" story. This is captured by the interaction term between scale and income. For the spillover threat, if a locality is surrounded by other localities with a high per cent of poor, we might expect the locality, as it gets richer, to try to fend off local migrants by withholding services. We add in the share of the population in the rest of the urban area that is poor, in this case in the bottom 20 per cent of income in the national urban sample of localities, and that variable interacted with own median income. We would expect richer communities to try to exclude the poor in the rest of the urban area.

Our empirical strategy is to estimate eq (3), with these basic income, size, and strategic behavior variables, adding in controls, *X(t-1)*, that seem relevant. We have three sets of controls. The first represents first and second nature geographic variables. What we like to know is water table and soil condition information, but lacking these on a national basis we try other controls.

We control for latitude (temperature) and altitude (moving inland to the inland plain). We also control for urban density, which reduces the per-household cost of laying physical infrastructure and generates a poorer health environment.

The second set represents taste and institutions, where we have a dummy for the south and southeast regions, which even after accounting for income differences may have a history/taste for better institutions compared to other regions. We also have the share of the population voting for pro-democracy parties in the 1982 national elections to the congress, representing a general "taste" measure in national elections for improving conditions for the poor, not based on local strategic considerations. Third we want to control for cost factors relating to the strain on the infrastructure system induced by fast growth. We can control for the 1970-1980 household growth rate of the locality. We also have a contemporaneous economic shock measure. Positive economic shocks induce more in-migration, straining the public system. Economic shock is the sum of the increase in national urban employment, excluding the own urban agglomeration, in each "traded urban good" sector[1], multiplied by the share in local employment (adding in agriculture) of each sector. That shock might weaken as transport costs to the economic centre of Brazil, São Paulo, decline.

There are two other types of controls. First we can add urban area dummy variables to control more completely for regional economic and institutional effects, beyond simple latitude, altitude and a south dummy. This then gives identification based on intra-urban area differences in conditions. Second we can introduce the lagged dependent variable. That would better represent the accumulated strain and capacity issues to do with back-logged provision. But it also captures some of the influences of all locality specific unobservables that persist over time—things to do with unmeasured locality geographic and taste considerations not already controlled for.

<u>Identification.</u> Our focus is on recovering parameters reflecting strategic exclusionary behavior based on the terms $y(t-1)), \ln(\mathrm{Z}(t-1)), \ln(\mathrm{Z}(t-1))*\ln(\mathrm{y}(t-1)$ in equation (3). Identification is tricky. There are a variety of institutional, geographic, and taste variables which affect provision of services and could potentially also have an influence on $y(t-1))$ and $\ln(\mathrm{Z}(t-1))$. Hopefully our controls, and ultimately urban area fixed effects deal with this issue, so that any remaining influences are orthogonal to $y(t-1))$ and $\ln(\mathrm{Z}(t-1))$. However there could be, say, unobserved *locality* specific strategic elements, triggered by the specific local history of politics and institutions not captured by the simple $y(t-1)), \ln(\mathrm{Z}(t-1))$ specification and controls, with or without urban area fixed effects. Such influences would have affected provision in the past and as such could affect the $y(t-1)), \ln(\mathrm{Z}(t-1))$ covariates, biasing their coefficient estimates. Given cross-section estimation for 1991 with 1980 covariates (see below), we could try IV estimation. However the only instruments we have for these 1980 covariates are lagged values of covariates or related variables from 1970. They would only be valid if strategic exclusionary behavior didn't start until 1970, which seems unlikely. We discuss IV estimation later.

[1] Manufacturing: food and beverage, tobacco, textile, clothing, leather, wood, paper, publishing, coal and oil, chemicals and pharmaceuticals, rubber, non-metal goods, metal goods, machinery, electrical goods, transport goods, furniture, other manufactured goods. Services: transport, commerce, finance and insurance, construction, accounting and legal, architectural and engineering and related technical, medical and dental and agronomy, publishing and design, other services.

A more specific issue concerns a control for the lagged dependent variable. While this is another way to capture past strategic influences, its presence serves more than one purpose. If we want a causal interpretation as to the role of the lagged dependent variable, it would be to capture pure back-log effects. To capture this we would need to instrument for the variable, with variables free of unobserved influences such as the specific local history of politics and institutions. An instrument for this might be the 1970-1980 economic shock experienced by the city which affects 1980 service levels but is arguably exogenous to unobserved local conditions.

Sample Choice For sample choice, we have an overtime choice and a cross-section choice. For the overtime, the issue is that the behaviours and institutional framework governing public service provision has probably changed from before to after democratisation in 1988. We thus focus on service provision in 1991, reflecting choices made in the pre-democracy era, when strategic behavior by elites may have been more in force. For the cross-section choice, we need to pick a relevant sample of localities. There are three issues. Our sample includes localities that are defined as part of urban areas in 2000, but in the recent past were rural areas. To maintain a look at just predominately urban environments where central provision of sewer and water is relevant, we exclude localities that in the base period (t-1) are under 50 per cent urban. Second, in some small localities, there are few houses that meet our criteria of being small and estimates of frequency of servicing are thus noisy. We exclude those localities where the actual number of houses surveyed that meet our criteria are fewer than 10. Finally, we report results for the sub-sample of localities that are part of multi-locality urban areas. This eliminates 64 localities that are single locality urban areas. This elimination has no impact on results concerning locality income and size; and allows us to rely on urban area fixed effect results and estimate the impact of conditions in surrounding localities within an urban area on own locality behavior, something that is relevant for most of the sample.

3.1 Results on services

Table 4 shows the basic results on service provision. For water service to small houses in 1991, column 1 gives results without the control for a lagged dependent variable and column 2 adds in urban area fixed effects as well as the listed controls. Columns 3 and 4 repeat the OLS and urban area fixed effect specifications but now control for the lagged dependent variables. In general, the effect of controlling for the lagged dependent variable is to weaken the 70-80 growth rate variable and the controls for latitude, altitude and south dummy. The past value of services has a positive sign as expected. While we would like this to represent a capacity-catch-up issue, the variable clearly could capture omitted variables whose influence persists over time. Controlling for urban area fixed effects tends to sharpen results on income and strategic variables, especially the influence of the percent poor in the rest of the urban area. Columns 5 and 6 give results for full service provision, just for the lagged dependent variable specification, with and without urban area fixed effects. We focus on column 4 and 6 results where we control for both the lagged dependent variable and urban area fixed effects.

Controls. We start by discussing the results on controls, focusing on column 4, given results in column 6 are similar. As density increases, that raises servicing, both a cost and externality consideration. A one standard deviation increase in density raises servicing by 4 per cent (i.e., 4 per cent points), from a mean of 78 per cent. Share pro-democracy at the national level positively influences local provision, where a one standard deviation increase in vote share raises servicing by 3 per cent. Finally high growth and positive economic shocks strain the system, reducing provision (controlling for income). A one standard deviation in the local shock variable reduces servicing by 3 per cent. Finally in column 3 (not reported), altitude has a positive significant (at 10 per cent level) effect on servicing; but, once lagged servicing is controlled for, latitude and the south dummy have no effect.

Strategic elements. In terms of income, scale and strategic considerations, results for income and scale are remarkably consistent across specifications in the table, although statistically weaker for full services. Starting at low levels for both water and full service provision to small houses, increases in median income and numbers of households are both associated with service increases to small houses, representing both positive income demand and scale effects. However, at about average median income (9.5), increases in numbers of households then become associated with decreases in services, with the rate of marginal decline associated with population increases rising as income rises. This is the "superstar" story: richer communities act to exclude the poor as their size gets big.[2] For water, at 1.5 standard deviations above mean median income (9.5), a one standard deviation increase in size (1.4) leads to a 4 per cent reduction in water servicing.

Columns 4 and 6 show the even greater impact on own locality provision of services of the fraction of households in the rest of the metro area that are in the bottom 20 per cent of income in the national urban sample, based on intra urban area variation. This effect is fairly strong statistically for both water and full service provision. Once a locality's income rises above the average median income, increases in the per cent poor in the rest of the metro area lead a locality to reduce servicing of small houses. For water, a one standard deviation (.138) increase in per cent poor in rest of urban area, leads service to small houses in the own community to decline by 10 per cent, noting we are controlling for urban area fixed effects. The magnitude for full services is almost the same.

Counterfactuals. A worry in interpreting the results in Table 4 as reflecting strategic decisions of communities to try to exclude low income migrants is that they might simply describe underlying algorithms for all public service provision—describing when communities as a whole are likely to have more or fewer public services, absent discrimination across house types. To explore this possibility we estimated the same models for high quality houses (seven to nine rooms in 1991, six to seven rooms in 1980). Results are in Table 5.

In Table 5, in columns 1 and 2, we re-estimate the models in columns 4 and 6 of Table 4, reporting just the results on strategic, income and scale effects. In column 1 for water, all these variables are now insignificant. For full servicing in column 2 of Table 5, results on income and scale appear to be similar to those in column 6 of Table 4; however the key per cent poor in rest of urban area coefficients are insignificant. For the income-scale effects that are significant, the results differ from those in Table 4 in a critical way. The effect on locality size on provision is positive (based on point estimates) throughout almost the entire income range. Scale effects just tail off as income rises, with the effect only becomes negative for point estimates at 1.7 standard deviations of median income above the average. We conclude that the determinants of services to small and large houses differ markedly, with political-strategic effects applying to small houses.[3]

So far, we have looked at services in 1991 just after the end of the non-democratic era in Brazil. When we go to 2000 services, two things happen. First, by 2000, water is virtually universal in all our localities. Second, as part of the democratic era, Brazil has engaged in strong

---

[2] If we divide the sample into quintiles by size and income and interact these to create 24 dummy variables, it is in the top income and top two size quintiles, where the decline is pronounced for water.
[3] In fact when we break size and income variables into quintiles and interact to get 24 dummy variables relative to the base case, no effects are significant for either type of service.

efforts to extend services to all localities and citizens, engaging in efforts to upgrade and regularise neighborhoods, and correspondingly the ability of elites to withhold services is greatly diminished. In Table 5, columns 3 and 4 we report on water and full servicing of small houses. The water results are completely insignificant; the full service results are much weaker than in Table 4.

Instrumental variables. In terms of IV work, in Table 4, if we treat just the lagged dependent variable as endogenous and instrument with the locality economic shock for 1970-1980, the coefficient on the lagged dependent variable generally and surprisingly rises. But once we are worried about endogeneity, presumably income and size variables are at risk as well, for omitted strategic behavior considerations. Instrumenting with lagged values of variables from 1970 (for 1980 covariates) has little effect on results, especially on income, scale and strategic effects. Either endogeneity of covariates (other than lagged provision) is not an issue or instruments from just 10 years before are almost as contaminated as the covariates.

3.2 Lot size and parcel regulations [Section and Table 6 omitted for now]

4. Effect of service provision of locality growth and population composition

There are two ways the literature examines the effect of regulation on housing and population supply to a city. First is explicitly within the context of a housing market framework where researchers estimate the effect of regulation on a housing supply function, usually arguing that regulation lowers the price elasticity of supply. We don't have data on housing prices, nor does housing start to conduct that type of analysis. The second way extends the analysis to look at the effect on population change across metro areas of regulation, arguing that regulation makes housing more expensive in a locality and hence retards migration to that locality. We follow that general approach. In the housing literature one way to examine the issue (Glaeser et al., 200) is to calculate "exogenous" employment shocks to each area induced by changes in national production patterns, and see if regulations dampen or offset such shocks. The other way is to look at the effects of regulations more in an urban growth context, where the non-regulation covariates are city growth determinants, from urban growth models (Glaeser, Shleifer, Scheinkman, 1995, Black and Henderson, 1999 and 2003 and Au and Henderson, 2006). Such arguments arise out of our conceptual framework, based on equation (2) where $N_m = N\left(N_0,\ A,\ c_I, N_m^{NS}\right)$. We use this growth framework adding in the measured shock component as well. Other adjustments will be made to recognize that growth work has focused on urban area growth, rather than locality growth as here.

The basic estimating equation examines the growth rate of urban households in localities as a function of variables which affect productivity and cost of housing, as well as the quality of public infrastructure.

$$d\ln(N_{i,t}) = \beta X_{i,t-1} + \gamma R_{i,t-1} + \varepsilon_{it}\ . \qquad (4)$$

In (4) $X_{i,t-1}$ is a vector of locality attributes which affect wages and the cost of living. $R_{i,t-1}$ is the "regulatory" variable, describing the quality of servicing of housing units likely to be bought by migrants. In the standard framework, wages grow through local knowledge accumulation and technological development in the urban area labor market. Local knowledge accumulation is measured by the educational level within the urban area and acts to grow wages and productivity, inducing immigration to the urban labour market. Second, there are demand shifters which raise local output prices and hence wages, thus inducing in-migration. One such consideration is the employment shock experienced by the locality (see above). Then there are the housing supply conditions within the locality: ceteris paribus, supply for migrants shifts out if there is more land. It shifts back if the pre-existing population is larger or has higher education and income (within the locality, as opposed to urban area). The former implies less land left for development, as does the latter, since the pre-existing population will consume more space as income rises.

The key issue in estimation concerns the error structure. The growth literature 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 differenced out variables that might affect overall long run size, such as access to markets and institutional development, as they might vary across space and change slowly over time. As such in the literature one standard approach is to rely on OLS estimation of at least cross-sectional growth equations.

However, it seems likely that there are omitted variables which affect growth and persist 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. As we saw above, service supply is affected by past city demand pressures which strain the system, as well as strategic considerations. Strategic undersupply of public servicing of housing should reduce population growth. However “undersupply” also occurs due to unmeasured past demand pressures which led to strains on the system and delay in supply provision. Thus a low quality of supply may represent unmeasured good growth conditions which may persist to some extent over time. Such influences will bias the estimate coefficient upwards, understating the negative effects of regulation. To deal with this problem, we will try to instrument for the $R_{i,t-1}$, as well as potentially the $X_{i,t-1}$ if we believe they also are affected by persistent unobservables. Before discussing instruments, we need to define our various outcome measures and the conceptual issues at stake.

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 its modern economy. By 1991 Brazil is 75 per cent urbanized. The axis of industrialisation that had focused on São Paulo and Rio expands with substantial and on-going industrialisation of hinterland cities. As we saw in Figure 1a, today urban areas everywhere have about the same growth rate and if anything growth is more rapid in the north compared to the traditional region of growth—the southeast (de Mata et al., 2003). This change in economic regimes will be part of an identification strategy. Despite the shift and high level of 1991 urbanization, there remains on-going migration from the north to the south, as well as national and urban population growth. In our sample the number of urban households grows by 40 per cent from 1991-2000. While we look for growth effects from 1991-2000, we note this is the democracy era, when local and national policy makers engage in planning exercises and initiatives to extend urban services throughout neighborhoods of cities and to regularise *loteamentos* and *favela* settlements. However, we will look to see if implicit regulation through poor servicing of small houses in 1991, as studied in Section 3, affects growth from 1991-2000.

Instruments. We need to instrument for two types of variables. First are the implicit regulatory variables, or the quality of public services to small houses in 1991. Second are growth drivers. From Table 4, political-institutional conditions and geography seem to affect quality of services but not modern growth. These include the 1982 national election vote results for the locality as a measure of local “tastes,” where a higher per cent pro-democracy vote is associated with better servicing. Second, the south and southeast of Brazil has traditionally had more developed institutions and modernized public sectors with an urban focus. And from Section 3, other things being equal, altitude is associated with better services.

More generally, we want to instrument for growth drivers such as educational attainment in the urban area, past size of the locality, and even economic shocks (which are based upon the base period, potentially endogenous industrial composition), These are all potentially influenced by variables which still affect locality growth. The general strategy is to use as instruments historical variables, which influenced urban growth in the past during the 1960-1980 period of rapid industrialisation and urbanization. These variables no longer drive growth today and the unobservables that influenced their attainment in the past do not persist today for two reasons as noted earlier: (1) The drivers of current national economic growth and hence 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. As a consequence, the places leading growth have changed as well. (2) Urban conditions have altered dramatically, due to economic growth, change in national output composition, and decentralization of heavy industry from central localities to hinterland areas. As a consequence, the “unobservables” driving local growth have changed. But historical variables remain strong instruments for current conditions, due to accumulation factors. If in the 1960s a locality attracted low education migrants versus high education ones who settled in the locality, that influences current educational composition, even if locality economic conditions have changed completely.

Use of historical instruments faces issues: (1) There is a tension between going further back in time to break the persistence in relevant unobservables and weakening the strength of instruments. (2) Apart from specification tests, it is difficult to “prove” that the assumptions are correct—i.e., this is not 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, a reason why we focus on the latter time period. For instruments, we use the following: (1) Access of a locality to São Paulo markets which played a critical role historically, before the development of modern trans-national transport systems even though today it has no impact on growth;[4] (2) the 1970-1980 economic shock localities experienced where a good shock improved local economic conditions at that time; (3) the illiteracy rate among the adult population in the locality and the rest of the urban area in 1970, which influences through accumulation average educational attainment today; (4) the manufacturing to service ratio in the *rest* of the urban area in 1970 which helped urban area economic attainment at that 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.

<u>Outcome measures.</u> Our first outcome measure from the simple model will be the growth in overall locality size, as measured by number of urban households. This is meant to capture the impact of restrictions on in-migration, as well as influencing departure migration decisions to the same effect. However public service restriction aimed potentially at “illegal” settlements would seem likely to have compositional effects as well. First, poor servicing of small houses would

[4] 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).

discourage illegal settlements and growth in numbers of small houses, relative to other types of houses. Second, this should influence the composition of the population.

However, sorting out influences on population composition turns out to be tricky. We might think the main impact would be on low skill households from the conceptual model. However there is the externality issue, as well as the issue of the servicing of larger houses for higher skill people, per se. While in principle, higher skill households may be strategically trying to exclude low skill households, high skill households may also be adverse to the negative externalities from under-servicing of parts of the locality and higher skill migrants may avoid communities with poor servicing in general.

We will start by looking at overall growth of households. Then we examine the effect of regulations on growth of small houses, and growth of three skill groups based on educational attainment of the family head: those who have not completed primary school, those who have completed primary school but not high school, and those who have completed high school or more. For our urban households in 2000, this is a 50, 30, 20 per cent breakdown of household heads.

4.1 Effects of Servicing on growth in urban households.

Table 7 contains two sets of results, one for water servicing and the other for full servicing. The service variable has been redefined relative to section 3: it is now the percent of small houses **not** served, or a measure of poor, or restrictive servicing. Underservicing in the two dimensions is strongly correlated. For each type of servicing, we present two estimations of the growth model: an OLS and an IV version done by 2SLS, both with urban area clustered errors. We start by noting basic growth effects.

Growth drivers. Employment shocks to the locality and urban area education levels (affecting urban area labor markets and knowledge development) both positively affect urban area growth. While increases in land are of the locality weakly promote growth, increase in the base population and level of locality education both reduce it, through presumably basic crowding effects. Having more people raises land prices, as does increased consumption with higher education and incomes. Finally having a greater proportion of the locality population be rural is a population supply factor increasing locality urban size. The IV coefficients on base period locality education, number of urban households and economic shock all approximately double in absolute magnitude, relative to OLS estimates, while the effect of having more rural households becomes insignificant. Based on IV results, one standard deviation increases in urban area education and economic shocks lead respectively to 3 per cent and 17 per cent increases in decade growth. One standard deviation increases in locality education and number of urban households lead respectively to 9 per cent and 13 per cent declines in decade growth rates, which average 40 per cent.

Servicing. In both models servicing of small houses has no effect on overall growth. However, we note the OLS coefficients estimates are positive, while the IV ones are negative, reflecting the hypothesised direction of bias. To explore the effect of servicing more we turn to composition effects.

4.2 Composition Effects

In a first pass, at identifying compositional effects, we simply take the covariates from the overall growth model and apply them to the different sectors—growth of small houses and growth of the three education groups—with appropriate adjustments. In particular the control covariates include urban area average education and locality economic shock as in Table 7. Then locality average education, base period number of urban households, and land are replaced with the new base period level of the corresponding growth variable (for three of the four specifications this is an education variable) and density (having land and base period size produces similar (insignificant) results).

In Table 8, we report the coefficients on just implicit regulatory variables. We report OLS and IV estimates, which all show expected bias in the same direction. There is a fairly compelling and intriguing pattern to the IV results. Low quality servicing of small houses reduces the growth rate of small houses. One standard deviation increases in under-servicing for water and for full service respectively reduce the growth of small houses by 8 per cent and 16 per cent (where the mean growth rate for the decade is 16 per cent). However that does not deter the growth rate of low education households. While OLS results show positive effects from poor service, the expected bias, IV results show no effect. However the results show strong negative effects on the growth of middle and high education households. Consider full services. A one standard deviation increase in poor servicing leads to a 18 per cent decline in the growth rate of middle education households (which declined on average by 28 per cent) and a 10 per cent decline in growth rate of high education households (which grew by 69 per cent in the decade). This could either suggest strong negative externality effects on the more educated from poor servicing of the informal sector in localities, or that poor servicing of small houses is strongly correlated with poor servicing of houses for more educated people, and that is an omitted variable.

We experimented with adding the servicing of large houses to the equations as a separate covariate. However that effect in levels is difficult to identify, especially in IV work with the instruments we currently have on hand. Our approach is to difference what would be the separate equations for growth of low (L) and higher (H) education households: $d\ln(N^k{}_{i,t}) = \beta^k \ln X^k{}_{i,t-1} + \gamma^k \ln R^k{}_{i,t-1} + \varepsilon^k{}_{it}, \quad k = L, H$. This has the advantage of differencing out location observables and unobservables whose effects are common to both groups. These would include geographic factors affecting the costs of services, lowering provision to all. As such, based on experimentation, we drop the average education variables and the economic shock variable from Table 7. Our estimating equation is

$$d\ln(N^L_{i,t}) - d\ln(N^H_{i,t}) = \beta_0 \ln(N^L_{i,t-1} / N^H_{i,t-1}) + \beta_1 \ln density_{i,t-1} + \gamma \ln(R^L_{i,t-1} / R^H_{i,t-1}) + \varepsilon_{it} \,. \quad (5)$$

Results are in Table 9. In Table 9, as the ratio of lack of servicing for small relative to large houses rises, that significantly lowers the growth in the ratio of low to higher education households, the basic exclusion result. IV effects are much larger than OLS ones, the expected direction of bias from earlier discussion. A 1 per cent increase in the ratio of relative poor servicing leads to a .2 per cent decline in the growth rate of low to higher education households.

For the control for the base ratio of low to higher education households, there is the usual reversion to the mean, negative effect. Controlling for that ratio, an increase in density lowers the growth rate of low to higher education households. This last effect may seem odd, but likely reflects two things. Higher education households tend to live in more residential communities, with lower industrial composition. Also as we saw in earlier figures, they tend to live in high population localities (that in essence are filled up), with less agricultural activity.

In Table 10, we show the results extend to the 1980-1991 time period as well. Again, as the ratio of lack of servicing for small relative to large houses rises, that significantly lowers the growth in the ratio of low to higher education households, the basic exclusion result. Note the specification tests for the IV results for 1980-1991 deteriorate badly, reflecting difficulty in finding 1970 instruments that are orthogonal to error terms.

5. Conclusions (to follow)

REFERENCES

Ades, A. F. and E.L. Glaeser (1995), "Trade and Circuses: Explaining Urban Grants," Quarterly Journal ofEconomics, 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” paper prepared for Lincoln Institute of Land Policy.

Biderman, Ciro (2007), Regulation and Informal Settlements in Brazil: A Quasi-Experimental Approach,” Lincoln Land Institute mimeo.

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

____________________ (2003), "Urban Evolution in the USA," Journal of Economic Geography, 3, 343-373.

da Mata, D, U. Deichmann, J. V. Henderson, S. V. Lall, and H.G. Wang (2007) “Determinants of City Growth in Brazil,” *Journal of Urban Economics*, forthcoming

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?” Paper prepared for Lincoln Institute of Land Policy, May 15.

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

Gabaix, X. (1999), "Zipf's Law for Cities: An Explanation," Quarterly Journal of Economics, 739-767.

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

Glaeser, E.L., J. Gyourko and R. Saks (2005a) “Why Have Housing Prices Gone Up?” *American Economic Review*, 95(2), pp. 329-333.

Glaeser, E.L., J. Gyourko and R. Saks (2006) “Urban Growth and Housing Supply,” *Journal of Economic Geography*, 6, 71-89.

Gyourko, J., C.J. Mayer and T. Sinai (2006) “Superstar Cities” NBER Working Paper No. 12355

Henderson, J.V. and A. Kuncoro (1996), "Industrial Centralization in Indonesia," World Bank Economic Review 10, 513-40.

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

Quigley, J.M. and S. Raphael (2005) “Regulation and the High Cost of Housing in California,” *American Economic Review*, 95(2), pp. 323-329.

Wu, W, (2004) “Sources of Migrant Housing Disadvantage in Urban China,” *Environment and Planning A*, 36, 1285-1088.

Zhou, M. and G. Cai (2005), “Trapped in Neglected Corners of a Booming Metropolis,” in *Urban China in Comparative Perspective*, J.R. Logan and S.S. Fainstein (eds.) forthcoming.

Table 1. Spatial allocation of rich and poor

	1970	1980	2000
Share of households that are rich (in top 10% of national urban households by income)			
Rio de Janeiro and São Paulo	16.0	16.2	16.8
Suburbs: Rio and São Paulo	12.3 [4.8]	12.1 [5.5]	11.6 [6.6]
Other central cities (multi-MCA urban areas)	9.2	10.7	14.1
Other suburbs	4.2	4.9	5.2
Share of h.h's that are poor (in bottom 10%)			
Rio de Janeiro and São Paulo	5.0	4.9	7.2
Suburbs; Rio and São Paulo	5.0	5.8	9.3
Other central cities	14.0	11.8	9.5
Other suburbs	14.1	12.8	11.2
Ratio: share rich to poor			
Rio de Janeiro and São Paulo	3.2	3.3	2.3
Suburbs: Rio and São Paulo	2.5	2.1	1.2
Other central cities	.66	.91	1.5
Other suburbs	.30	.38	.46

Except for "[.]", table contains weighted averages.

Table 2. Servicing of housing in Brazil in urban areas
All urban households living in localities over 50 per cent urbanized and 10 per cent service levels.

a. **All housing: over time provision and 2000 breakdown**

	Per cent with central water connection	**Per cent with full service: electricity, water and central sewer**	**Share of housing**	**Number of localities Water [full service]**
1970	59	41		226 [149]
1980	81	53		365 [214]
1991	91	62		428 [250]
2000	92	64		435 [359]
1991 breakdown				
Suburban	87	59	34	
Own house& land	92	65	61	
Own house, not land	82	34	8.1	
rent	94	67	21	
Ceded, other	87	57	8.5	
Migrants: bottom 20%	81	43	3.9	
Non-migrants: bot. 20%	87	52	15	

Table 2. Servicing of housing in Brazil in urban areas (continued)
All urban households living in localities over 50 per cent urbanized and 10 per cent service levels.

b. Services by house quality

Urban housing in center cities							
	Percent with water				**Percent with full service**		
	1980	**1991**	**2000**		**1980**	**1991**	**2000**
Low quality housing: 1-2 rms. 1980; 1-3 rms. 2000	59	79	87		18	36	49
High quality housing: 6-7 rms. 1980; 7-9 rms. 2000	88	94	95		62	70	69
All urban housing in urban localities							
	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

Table 3. Use of local land use regulations

	Per cent 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
Passed parcel law by 1991	51 [64]	.78	1.12
Passed parcel law by 2000	82 [100]	.69	1.11
Passed land use (zoning) law by 1991	43 [.55]	.79	.99
Passed land use (zoning) law by 2000	73 [96]	.70	1.62

Table 4. Service results

	Per cent small houses with water, 1991				**% small houses with full service, 1991**	
	(1)	(2)	(3)	(4)	(5)	(6)
Ln (# h.h., '80)	.354* (.182)	.451** (.161)	.324** (.160)	.437** (.153)	.261* (.140)	.223* (.134)

Ln(income '80)	.607** (.186)	.719** (.194)	.413** (.161)	.653** (.197)	.498* (.150)	.409** (.173)
Ln(# hh. '80)* ln (income '80)	-.038** (.018)	-.047** (.014)	-.035** (.016)	-.046** (.016)	-.028* (.015)	-.023 (.014)
% rest of urban area poor '80	.873 (1.97)	7.91* (4.35)	1.01 (1.42)	7.19* (4.32)	2.08 (1.36)	7.29* (3.86)
% rest poor rest UA *ln(income)	-.066 (.203)	-.802* (.444)	-.092 (.148)	-.740* (.445)	-.251* (.145)	-.751* (.396)
Service to small houses '80	n.a.	n.a.	.413** (.062)	.245** (.043)	1.03** (.109)	.620** (.096)
Ln(pop. density in MCA, 1980)	.025** (.0092)	.044** (.010)	.025** (.0070)	.027** (.0091)	.014 (.011)	.033** (.014)
Share vote pro-democracy '82 (national)	.157 (.102)	.183** (.077)	.170** .078)	.196** (.075)	.101* (.059)	.409** (.173)
Growth rate in locality of h.h. 70-80	-.037)* (.186)	-.019 (.019)	.012 (.017)	-.0010 (.016)	-.013 (.015)	-.029* (.017)
Locality economic shock 80-91	-.175 (.073)	-.167** (.058)	-.062 (.055)	-.105* (.059)	-.079 (.052)	-.107** (.053)
Dummy south, latitude, altitude	yes	no	yes	no	yes	no
Urban area fixed effects	no	yes	no	yes	no	yes
N	284	283	283	283	283	283
adjR^2	.43	.67	.61	.71	.81	.86

Standard errors in parentheses. * significant at 10 per cent level; ** significant at 5 per cent level. Means (s.d.): ln(income '80): 9.5 (.38); ln(#hh.'s): 9.4 (1.3); per cent poor in rest of urban area: .21 (.14).

Table 5. Public service strategies for large houses in 1991 and small houses in 2000

	Large houses, 1991 ('80 covariates)			**Small houses 2000 ['91 covariates]**	
	Per cent with water	**Per cent with full service**		**Per cent with water**	**Per cent with full service**
	(1)	(2)		(3)	(4)
Ln (# h.h., '80 [91])	.062 (.125)	.470** (.182)		.026 (.066)	.082 (.068)
Ln(income '80 [91])	.094 (.171)	.582** (.227)		.067 (.114)	.209* (.124)
Ln(# hh. '80 [91])*	-.0062	-.046**		-.0028	-.012

ln (income ’80 [91])	(.013)	(.019)		(.0088)	(.0095)
Per cent rest of urban area poor ’80 [91]	2.09 (3.69)	5.97 (5.28)		-.834 (1.47)	.184 (1.56)
Per cent rest poor rest UA *ln(income)	-.206 (.386	-.609 (.540)		.084 (.205)	.0024 (.217)
Controls from columns 4 or 6, Table 4	yes	yes		yes	yes
N	296	297		356	356
adjR^2	.64	.43		.72	.87

Table 7. Growth of urban households

	Growth in urban households 1991-2000, water provision		**Growth in urban households 1991-2000, full service provision**	
	OLS	2SLS	OLS	2SLS
Urban area avg. education 1991	.038** (.017)	.044* (.023)	.039** (.017)	.041* (.023)
Locality avg. education 1991	-.054** (.014)	-.083** (.030)	-.057** (.014)	-.082** (.028)
Ln(no. urban h.h’s, 1991)	-.032** (.013)	-.085** (.043)	-.031** (.013)	-.081* (042)
Ln (land area)	.010 (.010)	.046 (.029)	.011 (.010)	.043 (.023)
Share h.h’s rural in locality, 1991	.419** (.145)	.073 (.439)	.403** (.148)	.275 (.356)
Economic shock to locality 1991-2000	.495** (.115)	.824** (.396)	.494** (.112)	.860** (.399)
Share small houses without water 1991	.066 (.050)	-.024 (.104)		
Share small houses without full			.046	-.071

service, 1991			(.047)	(.068)
N	365	364		364
Rsq	.26		.26	
Sargan statistic [p-test]		7.2 [.52]		8.0 [.44]
Minimum first stage partial F [partial Rsq] {Anderson-LR test p- value}		13.9 [.41] {.00}		25.8 [.41] {.00}

Instruments are: locality adult illiteracy rate 1970, adult illiteracy rate in rest of urban area 1970, ln(distance to São Paulo), Share non-military vote 1982, dummy for south, 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 São Paulo), adult illiteracy rate in rest of urban area 1970* ln(distance to São Paulo), ln (no. of urban h.h's in rest of urban area 1970), altitude, share h.h.'s that are rural in rest of urban area in 1970, urban area economic shock 1970-1980, and urban area economic shock 1970-80* ln(distance to São Paulo).

Table 8. Compositional effects of regulation.

	Share small houses without water, 1991:		**Share small houses without full service, 1991**	
	OLS	IV	OLS	IV
(1) Growth rate of small houses 1991-2000	.106 (.124)	-.432* (.252)	-.226** (.084)	-.541** (.140)
2) Growth rate of low educ. h.h's 91-00 (not finish primary school)	.129* (.073)	-.175 (.176)	.222** (.057)	.114 (.073)
(3) Growth rate of medium educ. h.h's 91-00 (not finish high school)	-.095 (.110)	-.760** (.298)	-.193** (.073)	-.607** (.151)
(4) Growth rate of high educ, h.h's 91-00	.116 (.080)	-.237 (.215)	-.038 (.063)	-.329** (.139)

In all specifications, the covariates are the same as in Table , except ln(land area) and ln(no. urban households 1991) have been removed and ln(density, 1991) and the lagged dependent variable have been added. Instruments are as in Table 7, where all variables are treated as endogenous except land area.

Table 9. Growth in number of low to higher education households, 1991-2000
[ln (# low educ h.h/# higher educ hh) (t) - ln (#low educ h.h/# higher educ hh) (t-1)]

	water				Full service		
	OLS	**UA Fixed effects**	**2SLS**		**OLS**	**UA Fixed effects**	**2SLS**
Ln (share small houses/ share large houses, without water, 1991):	-.060** (.024)	-.027 (.021)	-.219* (.113)				
Ln(share small houses/ share large houses, no full service '91):					-.112** (.036)	-.043* (.025)	-.190** (.069)
Ln (# low educ /# higher educ), '91	-.106 (.079)	.017 (.049)	-.092 (.078)		-.167** (.083)	-.016 (.050)	-.218** (.102)
Ln(density, '91)	-.023* (.014)	.012 (.014)	-.016 (.013)		-.027** (.013)	.0010 (.013)	-.0317** (.015)
N [clusters]	334 [58]	334 [58]	315 [58]		355 [58]	355 [58]	336 [58]
Adj. Rsq	.07	.43			.15	.37	
Sargan stat. [p-value]			6.9 [.74]				8.2 [.61]
Minimum 1st stage partial F [partial Rsq] {Anderson-LR stat, p-value} (Cragg-Donald F)			9.0 [.20] {00} (5.5)				25 [.52] {00} (21)

Instruments are: locality adult illiteracy rate 1970, adult illiteracy rate in rest of urban area 1970, share non-military vote 1982, south dummy, manufacturing to service employ ratio in rest of urban area 1970, ln (no. of urban h.h's in rest of urban area 1970), altitude, share h.h.'s that are rural in rest of urban area in 1970, economic shock 70-80, ln(land area), altitude, per cent poor in MCA 1970, Dummy if full service < 10 per cent, 1970.

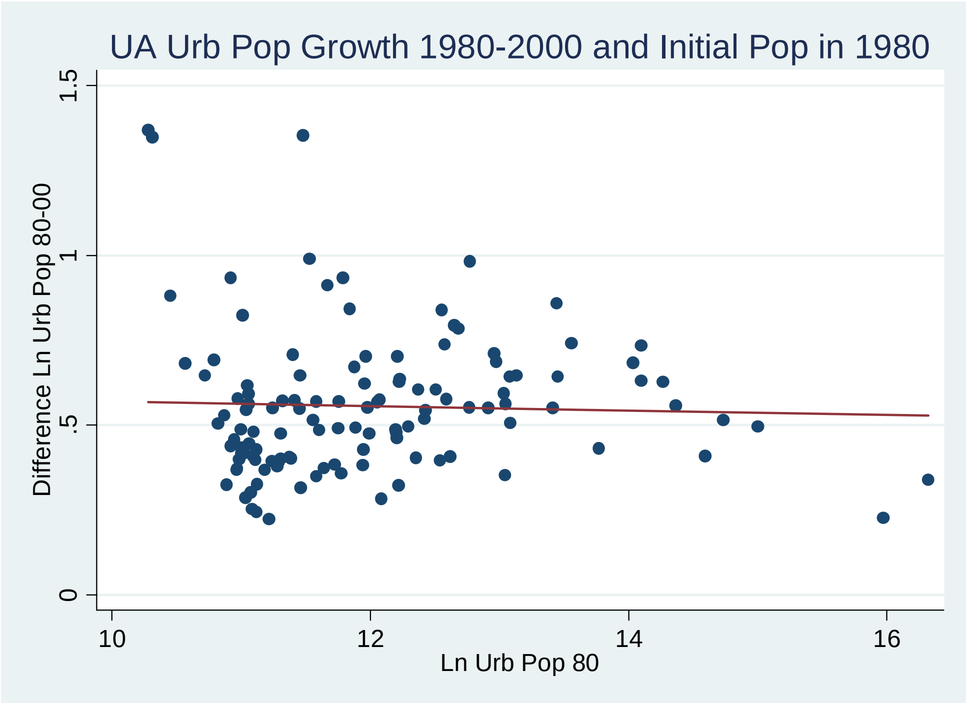
Table 10. Growth in number of low to higher education households 1980-1991
[ln (# low educ h.h/# higher educ hh) (t) - ln (#low educ h.h/# higher educ hh) (t-1)]

	water				**Full service**		
	OLS	**UA Fixed effects**	**2SLS**		**OLS**	**UA Fixed effects**	**2SLS**
Ln (share small houses/ share large houses, without water, 1980):	-.088** (.018)	-.091** (.020)	-.185** (.045)				
Ln(share small houses/share large houses, without full service '80):					-.067** (.019)	-.135** (.023)	-.083** (.028)
Ln (# low educ /# higher educ), '80	-.322** (.036)	-.338** (.077)	-.079** (.023)		-.343** (.032)	-.378** (.076)	-.334** (.102)
Ln(density, '80)	-.071* (.012)	-.093 (.012)	-.016 (.013)		-.076** (.012)	-.100** (.012)	-.031** (.022)
N [clusters]	291 [58]	291 [58]	276 [58]		297 [58]	297 [58]	282 [58]
Adj. Rsq	.45	.55			.44	.56	
Sargan stat. [p-value]			11.6 [.11]				14.2 [.047]
Minimum 1st stage partial F [partial Rsq] {Anderson-LR stat, p-value} (Cragg-Donald F)			9.5 [.26] {00} (8.1)				22 [.52] {00} (21)

Instruments are: locality adult illiteracy rate 1970, adult illiteracy rate in rest of urban area 1970, share non-military vote 1982, south dummy, manufacturing to service employ ratio in rest of urban area 1970, altitude, share h.h.'s that are rural in rest of urban area in 1970, per cent poor in MCA 1970, Dummy if full service < 10 per cent, 1970.

Figure 1. Urban area and MCA population growth 1980-2000

a) Urban areas [slope coefficient (standard error) of -.00644 (.017); R^2=0]



b) MCA's [slope coefficient (standard error) of -.0898 (.011); R^2=.14]

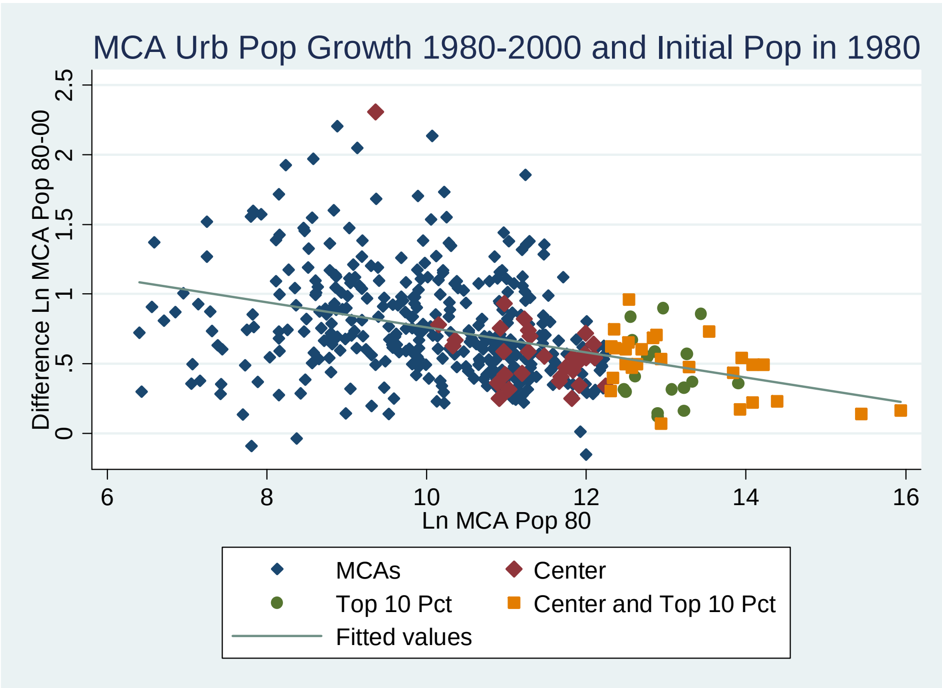
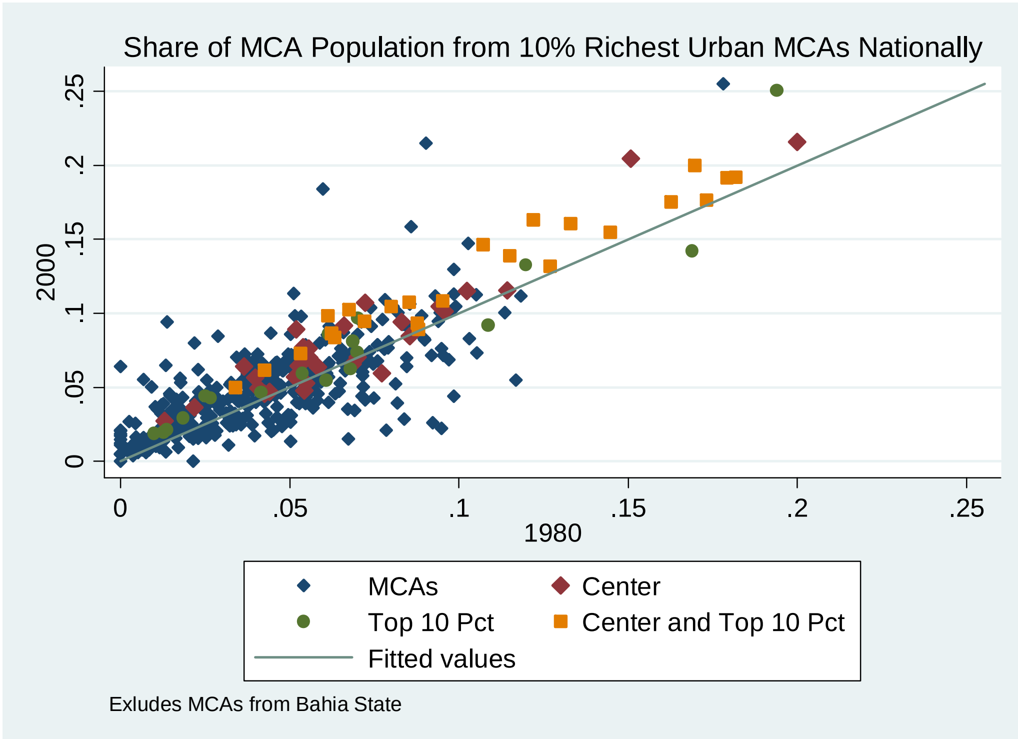
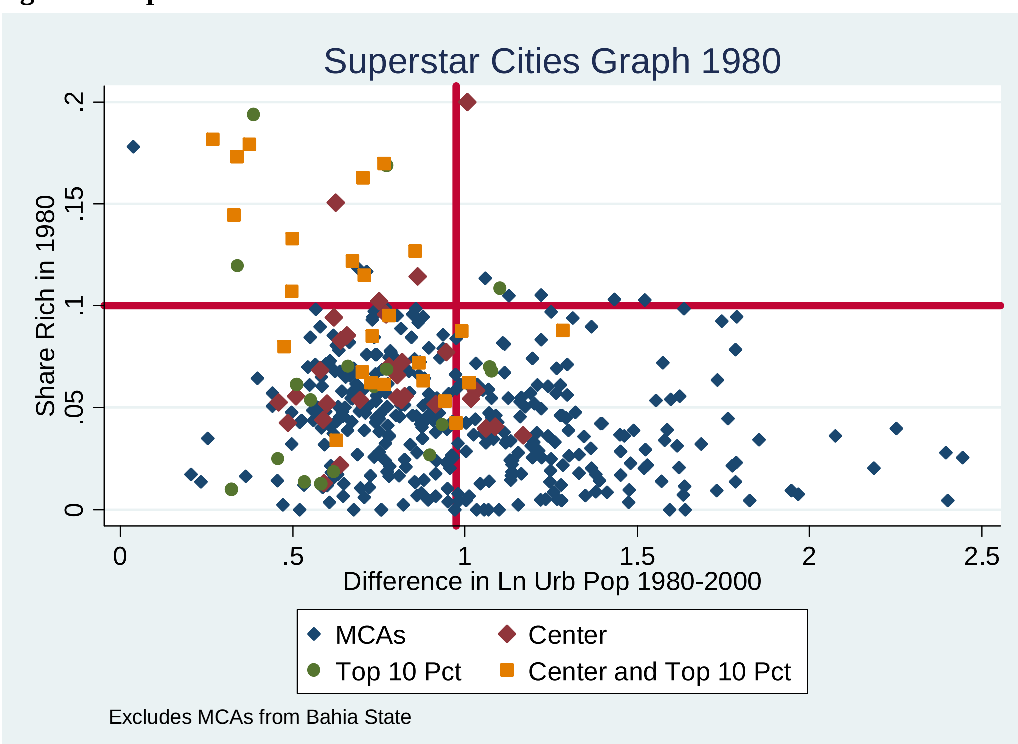


Figure 2. Share rich in 1980 versus 2000



Regression (s.e. in parens.): share rich 2000= .0113 (.00208) + .943 (.0365) share rich in 1980 + .112(.0356) share rich in 1980* dummy for center city. R^2 = .80. Sample is all MCA's over 15,000 urban population and 50 per cent urbanized in 1980; N= 287.

Figure 3. Superstar MCA's



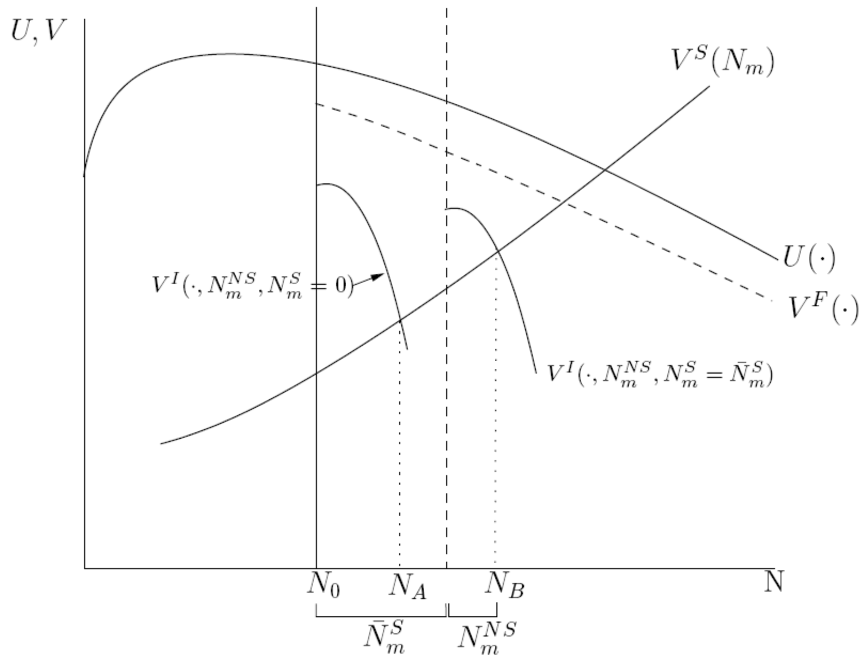


Figure 4: Equilibrium Level of Migration

APPENDIX, 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 rule for localities separately. We look at São Paulo and Rio *municpios*. We don't 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 *municipio*, 6 types of wall materials, 7 types of floors, and 8 types of roof, and whether the unit is single family residence. We then control for a variety of servicing features discussed momentarily. 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 minimise this problem, we insert district fixed effects district level fixed effects, where São Paulo has 56 and Rio 24 districts. The most recent year we can do this for is 1980—later years either don't have rent data or don't 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 2 shows the basic results.

In Table 2, the reported coefficients reflect the percent by which rents rise. From columns 1 and 3, it is clear in both São Paulo and Rio, 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 premium, even more so if it is metered (legal), indicating both reliable supply and higher (amperage) effective service. Garbage collection has modest or no impacts. Central sewer is much more valued than septic, especially in Rio. Septic only raises premiums modestly above having no service, presumably reflecting the failure of septic systems in these dense localities. Clearly there could be sub-district conditions that vary with these conditions, but the results are suggestive.

Table A1. Rent hedonics: ln (rent 1980)

	São Paulo			**Rio de Janeiro**		
	(1)	**(2)**	**Share (mean)**	**(3)**	**(4)**	**Share (mean)**

Water: other: inside plumbing	.121** (.030)		.0017	.161** (.040)		.0038
Water: well, inside plumbing	.185** (.011)		.023	.201** (.032)		.0067
Water: central connection, exterior	.121** (.0091)		.055	.229** (.019)		.0430
Water: central connection, interior	.285** (.020)		.897	.366** (.018)		.930
septic	.022** (.0054)		.308	.028** (.014)		.049
Central sewer	.094** (.0055)		.652	.177** (.012)		.916
Sanitation: for more than 1 house	-.0029 (.018)		.191	-.057** (.030)		.091
Sanitation: own house collection	.060** (.018)		.804	-.034 (.031)		.903
Electricity: no meter	.077** (.013)		.284	.192** (.033)		.162
Electricity: meter	.189** (.013)		.708	.342** (.034)		.833
Central water inside house		.181** (.0071)			.214** (.015)	
Full service (central water inside house, electricity, central sewer)		.195** (.0036)			.361** (.0074)	
Controls: house characteristics, district fixed effects	yes	yes		yes	yes	
N [districts]	196149 [56]	200,067 [56]		105135 [24]	108492 [24]	
Rsq	.65	.50		.58	.54	