Working Paper

OVE/WP-08/August 4, 2005

The Socio-Economic Impact of Favela-Bairro: What do the Data Say?

Fabio Soares and Yuri Soares*

Office of Evaluation and Oversight (OVE)

Inter-American Development Bank
Washington, D.C.
August 2005

For official use only

* Fabio Soares is with the Institute for Applied Economic Research (IPEA) and the International Poverty Centre (IPC). Yuri Soares is a Research Fellow at the Inter-American Development Bank (IADB) Office of Evaluation and Oversight (OVE). This evaluation represents collaboration between the IADB, IPEA, and the Rio de Janeiro Municipal Housing Office (SMH). The document has benefited from comments by Inder Ruprah, Kaizô Beltrão, Marcia Arieira, Tracy Betts, José Brakaz, Maria da Piedade Morais, Paulo Furtado, Suzanne Duryea and Guilherme Sedlacek. Data and technical advice from IBGE, Databrasil, Agrar, the Municipal Housing and Health Offices (SMH and SMS), State Health Office, IETS, and the Pereira Passos Institute (IPP) are gratefully acknowledged. This paper was also made possible by the excellent research collaboration from Renata Guagliardi and by the assistance from Dimitri Silva. The findings and interpretations of the authors do not necessarily represent the views of the Inter-American Development Bank, or their Executive Directors. The usual disclaimer applies. Corresponding authors: Fabio Soares and Yuri Soares, e-mail: yuris@iadb.org , Office of Evaluation and Oversight, Inter-American Development Bank, Stop B-750, 1300 New York Avenue, NW, Washington, D.C. 20577.

TABLE OF CONTENTS

INTRODUCTION

The first references to *favelas* in Brazil date back to the end of the nineteenth century, when soldiers of the Fourth Brazilian Expeditionary force, returning from a war against the separatist village of *Canudos* in the state of Bahia established residence in the hillsides of Rio de Janeiro, on the *Morro da Providencia*. Since this new pattern of urbanization resembled precisely the *Morro da Favela* settlement in *Canudos*, the name *favela* stuck and the *Morro da Providencia* became the first shantytown settlement in Rio. The idiosyncrasies surrounding the origins of the *favelas* notwithstanding, this type of informal, low-income settlement has over the decades continuously expanded to become, today, one of the city's most recognized and ubiquitous characteristics.

The *favela's* expansion and an increased segregation between formal city and the *favelas* have presented unique challenges for the formal city, as well as for the inhabitants of the *favelas*. From the point of view of the city the most obvious challenge is that *favelas'* establishment in public land represented a failure of property rights, producing a "tragedy of the commons". Lacking the appropriate price incentives, the pattern of informal urbanization holds little relation with the negative externalities imposed by their expansion, both in terms of environmental and health hazards, and in terms of a sub-optimal utilization of public land. On the other hand, the absence of state presence in the *favelas* and the uncertainty in tenure of informal settlements—particularly prior to the 1980s—has obvious welfare consequences to the inhabitants of *favelas*. Although there have been initiatives at the community level to fill the vacuum produced by the absence of the state, such as the establishment of neighborhood associations[1], most public goods, such as water, sewerage, electricity, and security remained underprovided. The consequences of this are evident in the high rates of vector-transmitted disease in informal settlements, the high incidence of geological disasters (mudslides, etc.), and the relatively low quality of housing. In the case of security, the absence of the state has opened up spaces for illicit entities, mostly related to drug trafficking, to take up functions related to policing and enforcement of norms[2].

Throughout this period of expansion, the nature of the relationship between the formal city and the *favelas* likewise changed, going from initial indifference, to rejection, to a more sympathetic and tolerant attitude. These changes in perception were accompanied by a dynamic response to the urbanization challenges mentioned above, with different administrations promoting policies that ranged from eradication and forced removal to urbanization and inclusion. The last of these policies was the *Favela-Bairro* project undertaken by the municipality of Rio de Janeiro in the mid 1990s. A clear break from earlier efforts to eradicate informal settlements, *Favela-Bairro's* expressed objective was that of urbanizing *favelas* and integrating them into the formal city. This document is an assessment of the *Favela-Bairro I* project. It is structured as follows: section I describes the public policy response to the problem of informal settlements in Rio de

[1] The nature of the functions exercised by the *associações de moradores* was determined in a idiosyncratic manner. Some of these associations simply handle mail for the *favela*dos, while others serve as a forum of discussion, dispute resolution, (informal) registry of property, and even (limited) provision of public goods from resident contributions.
[2] See Vianna (2004).

Janeiro from a historical point of view, placing favela-bairro in the context of this wider policy space. We also review the theoretical and empirical work surrounding the emergence and growth of informal settlements, and present a (mostly descriptive) characterization of the determinants of informal settlement in Rio de Janeiro. Section II describes the *Favela-Bairro* project. Section III describes the evaluation strategy, as well as the data and conceptual limitations. Section IV presents the results of the evaluation in (i) coverage of services (ii) earnings[3] (iii) conflict, violence, and health[4], and (iv) housing value and investment. Section V provides conclusions, policy implications, and discusses future research topics.

[3] In the case of employment, the incompatibility between the employment modules of the two census make comparisons extremely difficult.

[4] The source data for both health outcomes and conflict and violence are from causes of death. In this case one can identify death caused by diseases related to lack of sanitation, infant mortality and death by violent causes. Data on morbidity was available only for selected communities (participants of the *Programa Saude da Familia*) and only for 2000 onwards, making it impossible to assess the impact of *Favela-bairro* due to the lack of a baseline.

I. HISTORICAL AND THEORETICAL PRELIMINARIES

1.1 In this chapter we will describe the public housing policy pursued over the last few decades in Rio de Janeiro and describe the theoretical and empirical evidence on the emergence of informal settlements. This will serve a foundation for the results of the impact evaluation of *Favela-Bairro*.

A. Public Policy, Housing, and Informal Settlements in Rio de Janeiro [5]

1.2 The first observation to be made, widely documented in the literature, is that policy has been concentrated at the sub-national level; the Federal government has not had a large participation in financing housing solutions or in developing incentives for housing[6]. This policy vacuum at the top level has opened up space for states, and more recent municipalities to develop their own housing initiatives. This has certainly been the case with respect to low-income and informal settlements in Rio de Janeiro, which was originally dominated by state-level policy, and then, with the decentralization implemented through the new constitution of 1988, by municipal-level policy[7].

1.3 Prior to the 1950s society and state response to the question of *favelas* was only peripheral, always treated in the context of broader city code and planning strategies. They were also initially characterized by efforts to eradicate the settlements, with relocation a very distant secondary objective. The first example is the *Reforma Passos*, implemented from 1902 to 1906, which was responsible for building and infrastructure improvements in the center of Rio and in the emerging middle class boroughs in the center and south of the city. The same treatment was not accorded to informal settlements, which at the time were concentrated in the city center and were known as *cortiços, casas de cômodos and cabeças de porco*. During the *Reforma Passos* over 3,000 households in informal settlements were destroyed. Some of these residents were given houses in the city suburbs, but in most cases the only policy was of eradication. The emphasis on eradication continued unabated until the mid-1950s, with the *Plano Agache*, and the 1937 Building Code other prominent examples of policy—both at the rhetorical level and as part of initiatives—that essentially treated *favelas* as a impediments to urban development, provided no housing alternatives to those in informal settlements.

1.4 The first institution that legitimized the existence of *favelas*, by providing them with some measure of public goods services, was the Church. By funding the Fundação Leão XIII, the Church implemented the first neighborhood improvement project, providing urbanization and social services to *favelas*. This initiative serves as a catalyst for a change in perception regarding the *favelas*, which then engenders the creation of both a state and a federal agency charged with providing housing solutions for the poor. At the

[5] This section borrows heavily from Vianna (2004).
[6] The few housing policies that have been financed at the federal level (*carta de crédito*, BNH,) have in any case been concentrated in the middle class. For a review see World Bank (2002).
[7] The new constitution changed the transfer scheme, increasing the budgets of municipalities.

municipal level an agency to promote urbanization of *favelas* is created, the *Serviço Especial de Recuperação das Favelas e Habitações Anti-Higiênicas* (SERFHA), and three years later a similar agencies are created at the state and national level. The state and national initiatives had a somewhat different mandate of providing alternative housing solutions, and not one of urbanization[8].

1.5 These efforts are somewhat counterbalanced by an increase in the municipal eradication efforts, but nevertheless represent a clear break from the previous forty years of eradication efforts[9]. In addition, the increased role of the Church, through not only the Fundação Leão XIII, but later through the Cruzada São Sebastião, created in 1955, provided alternative mechanisms through which policy was articulated.

Table 1.1: Public Policy in Informal Settlements

Period	Type of Policy	Description
Prior to 1940	Eradication	**Passos Reform (1902-06)**. The reform had the objective or improving the center of Rio de Janeiro, eradicating informal settlements and *cortiços*. More than 3,000 households were destroyed; only 120 were built. **Plano Agache (1927)**. Innovative urbanization plan attempted to transform the southern neighborhoods of Rio into a "monumental city", providing new codes and building norms. Although the plan called for eradication, it did attest to the necessity of alternative housing solutions. The plan was not fully implemented due to the 1930 revolution. **The Building Code (1937)**. Favelas would not be part of the city's building code map. In an attempt to eliminate them the code prohibited any expansion of improvement of existing informal settlements, instead recommending the construction of the so called *parques proletarios* as housing alternatives for the
1940s	Eradication and Urbanization	**Parques Proletários (1941-42)**. The first *Parques Proletarios,* communities originally designed to be temporary housing while residents were being relocated, are built in Gávea, Leblon and Caju **Leão XIII Foundation (1946)**. Promoted by the Church, the foundation carries out some activities that improve living conditions for the favelas, marking the beginning of an urbanization policies outside the government context.
1950s to 1964	Alternating eradication and urbanization	**Cruzada São Sebastião (1955)**. This is the Church's second initiative in the *favelas*. Like the Leão XIII foundation, the initiative provides for urbanization improvements in favelas (12 favelas are improved), but it also takes on a role as intermediary between the state's eradication policies and the *favela* residents. **SERFHA (1956)**. First governmental agency created with the objective of urbanizing and improving favelas. **Federal Council on Housing (1960)**. Poor housing policies start to be implemented at the federal level, and start to attract attention at the national level. **COHAB (1960)**. At the state level, the popular housing company (COHAB) is created with the expressed objective of creating housing for the poor.

[8] At the state level the Companhia de Habitação Popular (COHAB) was created and at the national level the Conselho Federal de Habitação (CFH).

[9] During the period two commissions for the eradication of the *Favelas* were created, but they were never active in eradication efforts.

Table 1.1 (cont.): Public Policy in Informal Settlements

Period	Type of Policy	Description
1964 to1970s	Eradication	**Reformulation of the Leão XIII foundation (1963)**. The foundation is converted into a state agency. This represents an attempt to exercise control over the *favela* resident's association, which have gained political presence **CODESCO (1967) and CHISAM (1968)**.CODESCO represented the only initiative aimed at urbanizing the favelas during the military regime. It was, however, drowned by the creation of the CHISAM, which attempted to consolidate different housing initiatives, but in practice maintained eradication as a focal point. Between 1968 and 1975 roughly 100,000 houses would be eradicated.
1980s	Urbanization	**Promorar (1979)**. The program was implemented in order to urbanize informal settlements that were under flood conditions. It was managed by the National Housing Bank (BNH). Six *favelas* were intervened under Promorar. ***Projeto Mutirão* (1981)**. Implemented by the municipality, the project represented the first of a string of municipal-level interventions aimed at urbanization. Fifteen favelas, including the Roçinha, were partially urbanized. **Proface (1983)**. A state-level initiative that build on the municipality's Projeto Mutirão, it was focused on providing water and sewerage services to favelas. 60 favelas were linked to the city's water and sewerage. This was accompanied by a land tenure initiative, the *Cada Família Um Lote* program (1983), also implemented at the state level. **Programa Qüinqüenal (1985)**. The plan calls for incorporation of the favelas into the city more broadly than the provision of water and sewerage, recommending that all neighborhood state functions (formal recognition, mail, etc) be provided. It does not have an operative counterpart and is not implemented.
1990s	Urbanization	**Plano Diretor (1992)**. The new city plan incorporates the change in focus seen in the Programa Qüinqüenal, and due to new resources from municipal decentralization, it is able to fund broader neighborhood improvement projects. **Geap (1993)**. The newly created executive housing advisory board proposes six neighborhood improvement projects, including Favela-Bairro. **Favela-Bairro I (1994)**. Urban upgrading program implemented in 38 favelas. Improvements in other informal settlements are also provided for.
2000+	Urbanization	**Favela-Bairro II (2000)**. The second phase introduces activities in education, health and training, as well as community development. Property rights recognition are also included.

Source Vianna (2004)

1.6 Much of the gains of the late forties, fifties and early sixties were reversed with the 1964 military coup. The first military administrations, from 1964 to the late 1970s, accelerated the pace of eradication and curtailed, at the national level, all initiatives aimed at improving and urbanizing informal settlements. The efforts of the only urbanization initiative during the period, the CODESCO[10], were more than counterbalanced by the eradication efforts at the federal level, coordinated in Rio by the *Coordenação da Habitação de Interesse Social da Área Metropolitana do Grande Rio* (CHISAM). Between 1968 and 1975 roughly 100,000 households were eradicated. The military

[10] The CODESCO was in a way a precursor to the *Favela-Bairro*, since it presented a unified urbanization approach. Three *favelas* were improved during its tenure.

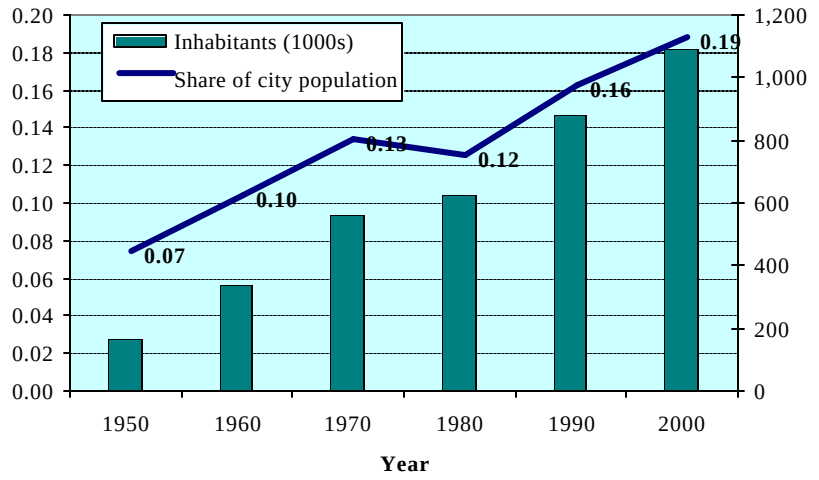
dictatorship also censored the emergence of community-level political and social groups, the *Associações de Bairros*, which had proliferated during the early sixties. This is done, ironically, by recognizing their legitimacy, imposing bureaucratic and legal restrictions on their activities, and providing conditional funding.

1.7 The eradication policies adopted during the dictatorship began to fragment and eventually disappeared in the last years of the military government (the years following the *Abertura*), and particularly throughout the 1990s, when *favelas* gained the status of legitimate, bonifide communities. The first sign of change came from the federal level, with the *Pro-Morar* initiative, aimed mostly at urbanizing six *favelas* that posed a high health hazard to the city due to their geographic location (they were located in flood-prone areas). At the municipal and state level two initiatives in the early nineties, the *Projeto Mutirão* and *PROFACE* were responsible for providing basic water and sanitation services to over 60 *favelas*, by hooking them up to the state water company. These efforts were complemented at the ideological level with new strategic formulations regarding urban planning. Initially in 1985 with the *Programa Quinqüenal*, a five-year plan for urban development, and later with the 1992 *Plano Diretor*, a more comprehensive urban planning framework, the city of Rio clearly established the legitimacy of the *favelas* and the need to incorporate them fully into the city's public services.

1.8 The formalization of this new policy on neighborhood improvement paved the way for six municipal housing initiatives undertaken during the decade, the most important of which was the *Favela-Bairro* project. One must note, however, that this new formulation was developed at the sub-national level. The federal response to housing needs had thus far been very marginal, and in the case of the *Banco Nacional de Habitação*, can only be regarded as a failure given that it neither addressed the housing needs of the poor, nor did it successfully recover resources lent to the middle class for new housing construction. In this sense, although the sub-national response to growing informality of housing among the poor explicitly recognized the needs of these residents, it had no response to the incentives that led to the establishment of informal settlements. In other words, the integration of the *favelas* into the city was not accompanied by a wider national policy of dealing with the future emergence of additional informal settlements.

1.9 The discussion above represents a broad brush approach to characterizing public housing policy for informal and poor housing in Rio. The lack of historical data on informal settlements and the composition of residents in these settlements over the last 50 years do not allow us to identify the impact of these policies, either in the standard of living of the *favelas* or even in the expansion of informal settlements more broadly. However, it is instructive, at least from an anecdotal point of view, to characterize the temporal evolution of the population in *favelas*.

Figure 1.1: Growth in Rio's Population and the share of *favelas*



Source: IBGE, Census, various years

1.10 Figure 1.1 presents the change in the *favelas'* share of the city population, along with the population living in *favelas*. As can be seen, the increase in informal settlements has been dramatic, increasing from 7% of inhabitants in 1950 to almost 20% by the end of the decade. There is one period in which we see a slight decrease in the percentage of city population residing in sub-normal units, which corresponds precisely to the harshest years of the dictatorship and the period with the most active eradication policies. This may have been responsible for slightly containing what would seem otherwise to have been a monotonically increasing informal share of city population. However, the fact that this decrease is only seen in one of the three periods of eradication suggests that the results of the hard-line removal policies implemented between 1950 and 1980 were overshadowed by more powerful incentives. In the next section we will describe alternative hypothesis that could explain the dynamics in the geographic distribution of housing in Rio de Janeiro, and relate this literature to the more recent *Favela-Bairro* initiative implemented during the mid 1990s.

B. Preliminaries on the Pattern of Informal Settlement in Rio de Janeiro

1.11 There is actually remarkably little in the way of research on the proliferation of informal settlements in dense metropolitan areas. Most of the research on the emergence of low-income communities has been in the developed world context, looking at the sorting of poor into ghettos, and therefore does not take into account the fundamental differences in institutions, property rights, and land markets between, say, Brazil and the Unites States. It is nevertheless instructive to review some of the literature addressing patterns of urbanization and their socio-economic consequences. Throughout the discussion two questions are paramount: (i) what are the incentives that determine the emergence and growth of *favelas* and (ii) what type of sorting mechanisms are responsible for the self-selection of residents into *favelas* versus other low-income housing neighborhoods.

1.12 The literature on geographic sorting in general focused on issues of efficiency and externalities of having communities segregated by income, race, skill, or other dimensions. This literature attempts to characterize the benefits and consequence of geographic sorting. In most cases the evidence shows that sorting results in worse socio-economic outcomes, ranging from schooling to crime and even faith[11]. The consequence of this literature from a policy perspective is that even if we do not have enough information to construct a geographical counterfactual to ethnic or income concentrated city, we do know that this pattern of urbanization does have significant impact on the quality of life of those who find themselves segregated. These questions do not, however, help to explain why people self-select into particular neighborhoods and what are the incentives and sorting mechanisms underlying these decisions.

1.13 A second strand in this literature in Urban Economics is concerned, precisely, with selection and sorting. This literature is derived from the seminal articles of Alonso (1964), Mills (1967) and Muth (1969)—or AMM for short—that addressed the post-war pattern of sub-urbanization observed in the United States. The model, in its simplest form, predicts that the wealthy will select farther locations if the income elasticity of demand for land is higher that the income elasticity of the cost of commuting, mostly determined by time of commute[12]. However, this model has been empirically debunked, given that efforts at estimating these elasticity parameters have not supported the model's predictions[13]. Given the empirical failings of AMM, competing hypothesis have emerged to explain the disparate patterns of urbanization.

1.14 The most recent development that explains the emergence of sorted neighborhoods rests on one of three foundations: (i) theories of differential amenities, (ii) theories of complementarity of human capital, and (iii) discrimination theories based on preferences. The amenities theories are straight forward. When public amenities are provided in the center the wealthy's demand for these amenities drives up the rent on land, which makes these districts prohibitively expensive for the poor. This theory is used to explain, for example, why European cities, characterized by high levels of central business center amenities, have a pattern of income concentration near the city centers, while most large American cities exhibit exactly the opposite[14]. The explanation based on human capital

[11] Borjas (1995) found that neighborhood ethnicity has an impact on socio-economic outcomes conditional on household demographics, social status, and income. Case and Katz (1991) look at the peer effects in Boston and find significant effect related to crime, alcohol and drug use, as well as school drop-out. Similar impacts are found by Cutler and Glaeser (1995) for African Americans living in segregated neighborhoods. Looking more specifically at crime, Glaeser, Sacerdote and Schneinkman find that wide regional variations of crime pockets are in part explained by neighborhood effects. The evidence that sorting produces negative outcomes is not, however, a consensus. See for example, Wilson (1987). He points out that sorting may serve to develop models in Black communities. Also, in some cases the negative impact of geographic sorting may be confounded with the negative attributes which give rise to self-selection in the first place. This argument is made, for example, by Cutler et. al (2002).

[12] The argument is of course somewhat of a tautological construct. For a review see Fujita (1989).

[13] In particular, research as early as 1977 (Wheaton) and as recent as 2000 (Gleaser, Kahn and Rappaport) find relatively modest income elasticity with respect to land area. Gleaser et all's estimates are at most 0.4.

[14] See Brueckner, Thisse and Zenou (1997). Also, Baum-Snow (2004) looks at the impact of expansion in road infrastructure on the pattern of decentralization in the United States and finds that much of this can be explained by reductions in the cost of commuting.

complemetarity, rests on the assumption that human capital benefits accrue to others in the form of lower costs of acquiring human capital. First proposed by Benabou (1993), this hypothesis leads to equilibria in which individuals migrate in order to obtain better schooling, leading segregation based on skill[15]. The last theory of sorting reviewed was originally proposed by Schelling (1972), in his "tipping point" formulation. Although the most *ad hoc* theory of those presented, since it presumes asymmetric intolerance of one group by another (say, intolerance of blacks by whites, for example) in any given neighborhood, it has gained considerable attention due to its simplicity and rhetorical appeal. The main result of this type of hypothesis is that sorting will occur, and that the end distribution of ethnic composition in equilibrium will be bimodal. Others have developed this model further, by looking at sustainability of equilibria, as well as looking at different distributions of innate individual taste parameters that can lead to different patterns of sorting[16].

1.15 The strand of literature above is related to the original question posed above: why have *favelas* proliferated, only to the extent that it sheds light on the fundamental factors that lead to housing-work decisions, the most important of them being the price of land, neighborhood amenities, and the opportunity cost of commuting. From the point of view of the theories assessed above, explanations based on opportunity costs, changes in public amenities and income elasticities are most attractive, given that these are all variables that have seen significant change in Rio de Janeiro over the last few decades. Explanations based on changes in tolerance and discrimination do not seem applicable in this setting given that perceptions surrounding *favelas* have in fact "improved" over time, a regularity reflected in the changes in public policy also. However, none of the explanations given above consider the additional complication associated with the informal nature of settlements and the perceived (and during many years real) threat of disappropriation of property due to lacking property rights. This is due to the fact that these theories have been concentrated around the developed world experience, in which property rights in land has not been a critical issue during the twentieth century. The role of property rights in urban settings has of course been looked at by many authors. But these studies tend to concentrate on the impact on investment decisions, property value and credit availability, derived from insecurity of claims on future assets and on the inability to engage in transactions of property, rather than on the dynamics of urbanization[17]. The link between these two separate issues is poorly developed in the literature. Among the few studies that attempt to address sorting and tenure insecurity is the study by Jimenez (1984), who estimates hedonic housing price functions as a function

[15] Benabou assumes that education is a club good: the cost of becoming a high-skill worker is decreasing with the proportion of the neighborhood that also becomes a high skill worker. In this framework, and without *ex ante* assumptions regarding ability or initial skill conditions, he shows that these externalities in the costs of education can lead to equilibria in which neighborhoods sort by skill, and that eventually lead to sorting by employment and income.

[16] See for example, Galster (1990), Cutler, Glaeser, Vidgor (1997), Möbius (2000).

[17] See for example Lanjouw and Levy (2002), Besley (1995), Olinto (2000), Field (2003), which all found favorable, impacts of titling on investment. De Soto (2000) also argues that increased titling allows people access to credit, although much research has found that this effect is relatively small or even undetectable. For a review see Woodruff (2001). One must, however, also note that authors have found that informal tenure agreements can be just as effective as formal agreements, as is pointed out by Kanazawa (1996) and Ellickson (1991).

of land tenure characteristics[18]. However, the Jimenez paper did not lead to significant additional work on the topic.

1.16 Moving from the broader issue of residential sorting to the much more specific issue of sorting into *favelas* in Rio de Janeiro, reveals that there is almost nothing done on the subject empirically in Brazil. Although there are a number of studies that describe the socio-economic profile of low income housing communities in the municipality of Rio de Janeiro, most using data from the *Pesquisa Socioeconomica de Comunicades de Renda Baixa* (a survey done in the late 1990s in select *favelas* of Rio de Janeiro), virtually none of these papers attempt to answer questions regarding the formation of *favelas* or even the decision to live in a *favela* versus other low-income options[19]. One exception is the work done by Morais, Cruz and **Oliveira** (2003). In their study the *favela* versus non-*favela* decision is modeled by looking at the IBGE's classification of a subnormal agglomerate census track as being synonymous with an informal dwelling. The main result that falls out of their estimation is that dwellers of informal settlements are more likely to be non-white, unemployed, migrant, and on average have shorter commute times[20].

C. An empirical assessment of the low income housing market in Rio de Janeiro

1.17 If hypothesis based on sorting are correct, and if residents in *favelas* are sorted on preferences for risk, on preferences of time usage, or on preferences regarding the availability of public amenities, one would expect some of these differences to be present in observables, to the extent that reduced forms of these observables are also functions of the unobservables on which residence is sorted. A starting point, then, for any empirical exercise of housing choice would be to describe the characteristics of *favela* and non-*favela* residents. Table 1.2 presents means for informal and formal residents in three different populations: (i) all of Brazil in the left panel; (ii) only urban residents in the second panel; and (iii) only low-income residents of the municipality of Rio in the rightmost panel[21]. The data are from stacked cross-sections of the *Pesquisa Nacional por Amostra de Domicilio* (PNAD), for the years of 1992-1995[22].

---

[18] As one would expect, his results are consistent with agents with higher risk characteristics residing in locations with lower quality tenure.

[19] Other studies have stacked cross sections of the Pesquisa Nacional por Amostra Domiciliar (PNAD).

[20] Two fundamental problems qualify these results, however. First, although the study contains dummy control for urban and rural areas, the two comparison groups (sub-normal agglomerates versus normal ones) are so disparate as to be almost meaningless. For the question we are posing a more sensible comparison would be sub-normal areas within the large municipality of Rio de Janeiro versus other poor communities. This reflects more appropriately a behavioral model in which a low-income resident can select to live in a informal dwelling within the metropolitan area versus other low-income options.

[21] This was done at the community level. That is, communities within the favela per capita income common support were included. One should note that "community" is defined by a census track. This may in fact not be the case since some *favelas* are composed of more than one track and since, in rare instances, a track will cover more than one *favela*. Results were not sensitive to the cutoff specification between one and two times the income level of favelas. Above two times the cutoff becomes relevant.

[22] This stacking is possible since each of the PNAD household surveys resample household from the same census tracks each year. The PNAD samples large municipalities with probability equal to one, and smaller municipalities with different probabilities, based on population. Within each of the municipalities, census tracks are randomly chosen and kept for the entire period between census years. From these census tracks households are randomly chosen each year, without replacement. The sample size is from 1992 to 1995. At this point we will abstract from

Table 1.2: Means and summary statistics for various types of urbanization groups

	Brazil	Large Municipalities [1]		Municipality of RJ		
					Low income	
		Formal city	Informal	Formal city	Formal city	Informal (Favelas)
Household head characteristics						
White	0.56	0.61	0.42	0.67	0.50	0.50
Years of education	6.1	7.7	5.1	8.5	6.2	5.1
Age	45.2	45.0	43.0	49.0	46.5	46.7
Single mother	0.16	0.20	0.19	0.23	0.20	0.19
Migrated	0.70	0.72	0.74	0.51	0.47	0.60
Proportion migrants from other state	0.36	0.42	0.43	0.41	0.34	0.50
Employed	0.78	0.75	0.77	0.69	0.71	0.74
Age composition of members						
Children (0 to 16)	1.60	1.30	1.75	0.92	1.21	1.19
Adults (17 to 64)	7.01	6.76	7.68	4.49	6.02	6.52
Elderly (more than 64)	0.72	0.83	0.30	0.43	0.26	0.27
Crowding (members per room)	0.84	0.79	1.16	0.69	0.83	0.87
Predicted rental rate	187.7	251.8	150.3	283.2	213.3	224.0
Household self-reported ownership	0.71	0.68	0.84	0.70	0.75	0.80
Number of rooms	5.5	5.5	4.3	5.5	4.7	4.6
Piped city water service	0.69	0.90	0.81	0.97	0.94	0.90
City sewerage service	0.38	0.60	0.34	0.75	0.58	0.76
Garbage collection	0.63	0.85	0.58	0.90	0.91	0.60
Commute time one hour of more	0.11	0.13	0.14	0.19	0.29	0.19
Household income per capita	277.93	464.80	152.44	551.48	231.17	194.04
Proportion of households						
Per capita income quartile 1	0.25	0.13	0.28	0.09	0.15	0.14
Per capita income quartile 2	0.25	0.19	0.32	0.17	0.26	0.28
Per capita income quartile 3	0.25	0.27	0.25	0.26	0.33	0.30
Per capita income quartile 4	0.25	0.41	0.14	0.48	0.24	0.26

[1] Source: PNAD. 1992-1995. Defined as municipalities of 100.000 or more inhabitants

1.18 Table 1.2 presents the summary statistics for *favelas* and non-*favelas*. The first observation is that the differences between these two groups vary depending on which population they are drawn from. For the sample of all large metropolitan areas (center panel), differences in commute time between formal and informal communities are negligible, 13 percent of formal communities commuting one or more hours to work, versus 14 percent for informal communities. This is completely reversed when one looks at the Municipality of Rio, where the difference jumps from 1 percent to 10 percent, for the comparison between *favelas* and other low income neighborhoods. In terms of household head demographics we see that overall, most household heads (70 percent) were either born in a different municipality or at some point lived in a different municipality. This number falls to just over 50 percent for the municipality of Rio de Janeiro. Also, when comparing migration patterns within the municipality and those of other large metropolitan areas, we see that the differences are much greater in Rio, with more migrants living in informal communities. We also see that in terms of migrants

problems arising from the possible non-stationarity of time series that may produce cointegration of variables that change together over time. However, given the short nature of the census track panel, this is not expected to be a problem; almost all variation will be from the longitudinal dimension of the data.

from out of state, there is a greater proportion among *favelas* than among other low-income Rio communities. With respect to the race profile, once again we see differences across panels. In the large city sample, informal settlements are more likely to be non-white; this racial asymmetry is, contrary to widely held belief, not true for poor settlements in Rio. Although there is a clear racial divide among poor and non-poor (as expected) this does not translate over to formal and informal settlements. These differences suggest that city idiosyncrasy is quite important in analyzing patterns of urbanization: not all informal settlements are *favelas*.

1.19 Concentrating on the right panel we see that the largest differences in Rio are between poor and non-poor communities and not between formal and not formal. In terms of race, schooling, as well as dwelling characteristics, poor communities, as expected, show depressed indicators relative to their richer counterparts. However, the variables that are most markedly different between *favelas* and non-*favelas*, are the distance to and from work (proportion of household commuting more than one hour), which is substantially lower for *favelas*. In terms of access to amenities, we see one characteristic that set *favelas* apart from the formal city. Although *favelas* have lower access to piped water, the availability of garbage collection shows the most dramatic gap, of 34 percentage points. The higher sewerage coverage in *favelas* reflects the fact that this measures only direct connections to the city. It does not take into account septic sewerages, which may be more common in the peri-urban areas—where *favelas* are not common. This is consistent with sorting on time preferences versus public service amenities, although the observed differences are certainly smaller than expected.

1.20 The last comment in order is with respect to the income distribution of those in formal and informal areas. Conventional wisdom dictates that *favelas* are very poor communities, with poor sanitary and public services characteristics. However, the table above shows a different reality. The bottom panels show the proportion of households in each income quartile. Since the quartiles are defined based on the country's income distribution, at the country level exactly 25 percent of the population is in each quartile. Moving to the rightmost panel, we see that the population of *favelas* is not poorer than the country as a whole. Although more concentrated in the middle quartiles, *favelas* have a slightly better distribution. Comparing *favelas* with the Municipality of Rio, we see that in this case it is true that they are much poorer; almost 50 percent of the city is in the top quartile. In other words, although these communities are certainly poor vis-à-vis the city, they are on average just as poor as the country as a whole. The distribution between *favelas* and other low income communities are also similar.

1.21 Of course the Table above only allows for pair wise comparisons, which cannot take into account the correlation among the different variables that may determine housing location. To account for this—and borrowing from Morais et al—we model the probability that a resident chooses to live in a *favela*, based on personal characteristics, including time-to-work. If the amenities theory holds, the low-income urbanization sorting in *favelas* is driven by a lack of public services, but may be partly offset by a greater proximity to work, as is suggested by the summary statistics presented above. For greater comparability, we have segmented the housing market, so that only poor

communities are included, where the income cutoff was arbitrarily established at the same income levels of *favelas*[23]. The empirical specification is as follows:

$$p(\textit{favela}) = \Phi(b_S X_S + b_H X_H + b_D X_D) \qquad (1.1)$$

where X_S denotes characteristics of the household head, such as age, schooling and ethnicity (white versus non-white), X_H denotes household characteristics, and X_D denotes travel distance to work. In addition to the variables described above, the probit equation also includes the average rental rate for each census track, denoted neighborhood rental rate in the table below[24].

Table 1.3: Probit model for housing choice in a *favela* (versus other low-income community)

Favela Probit	dP/dx	z-statistic	mean
Household head characteristics			
White	0.017	(1.05)	0.604
Years of education	-0.021	(-8.44)	8.138
Age	-0.002	(-1.89)	41.760
Single mother	0.016	(0.68)	0.146
Migrated	0.043	(1.61)	0.427
Migrated from other state	0.109	(3.87)	0.334
Age composition of members			
Children (0 to 16)	0.002	(0.40)	1.174
Adults (17 to 64)	0.001	(0.72)	4.004
Elderly (more than 64)	-0.077	(-3.34)	0.103
Crowding (members per room)	0.012	(0.64)	0.772
Neighborhood rental rate (1000s)	-2.766	(-17.20)	0.135
Household self-reported ownership	0.149	(7.60)	0.710
Number of rooms	-0.021	(-2.94)	5.046
Piped city water service	-0.063	(-1.82)	0.963
City sewerage service	0.282	(16.11)	0.700
Garbage collection	-0.489	(-23.00)	0.874
Commute time one hour of more	-0.138	(-7.48)	0.234
Household income p.c. (1000s)	0.054	(2.49)	0.235
Pseudo R-squared	0.271		
Observations	4597		

1.22 As can be seen from Table 1.3, the probit results largely corroborate what was seen in the summary statistics. Of particular note are the marked differences in education and out of state migration effects, as well as the different effects of amenities and the proximity to work: household heads that commute less than one hour are sixteen percentage points more likely to live in *favelas*. With respect to migration, we see that, conditional on

[23] See footnote 21.

[24] The neighborhood rental rate for observation *i* is based on all neighboring observations, but excludes the *i*th observation.

being a migrant, those who arrive from outside the state of Rio are almost eleven percentage points more likely to live in a *favela* than in other low income communities. In terms of amenities, the only variable that showed a conditional decrease in the probability of living in a *favela* is garbage collection. The presence of sewerage is associated with an increase in probability, mainly due to the fact that non-city sewerage systems are not tabulated as sewerage. However, beyond the direct measurement of amenities, the neighborhood also measures the quality of the neighborhood's housing. Here we see that even among low income housing, neighborhood high neighborhood housing prices are still associated with a lower probability of residing in the *favela*. For every 10 Reais in additional average neighborhood monthly rent, the probability of being in a *favela* decreases by a full 2.8 percentage points.

1.23 The probit model allows one to observe how characteristics endogenous to the household influence housing decisions. The results suggest that there is sorting. However, this sorting is not on basis of race, as would be expected by the "tipping" model of asymmetric preferences. Rather, the sorting is based on factors related to education and distance from work. Of course we cannot separate what part of this sorting is due to unobserved heterogeneity; nor does the approach allow for the estimation of relationships between community-level characteristics in *favelas* and in other communities, such as the availability of public services. The negative coefficient both on age and on time of migration suggests that there is also sorting over the lifecycle. In particular, it is possible that *favelas* may represent a temporary housing solution for households in transit. Without longitudinal information on households, however, it is not possible to test this hypothesis.

1.24 The conclusion that falls out of this analysis is that if sorting is based on distance-to-work (as an amenity) versus lack of public services, or even, say, lack of security (as a negative amenity), this has important consequences to the extent that programs such as *Favela-Bairro*, by providing public service amenities, may be changing the incentive structure that is driving the selection of residents into formal communities or into existing (or newly occupied) informal communities. By altering the incentive structure one of the first effects of upgrading neighborhood programmes can be a stimulus to migration towards the recently improved areas or into areas where expectations exist of future improvements. Beyond the market failures that this entails the program could also provide incentives for residents to migrate to areas that already have huge concentration of households. For this reason, one of the outcomes we will be looking at in this evaluation is precisely changes in population size before and after the programme. One should note however, that the expected direction of change in this case is ambiguous: neighborhood improvement could have a larger impact on informal settlement more broadly, and if property values are impacted, could actually hinder immigration.

1.25 This higher price could be the consequence of an increase in demand for recently improved communities, which in turn could change both the turnover of residents (housing tenure) and eventually the composition of residents. Residents in benefited areas could be well-off in terms of both wealth and current income. A shortcoming that can prevent this process from happening is the fact that *favelas* residents "choose" to live

there for its low costs[25] vis-à-vis other low income area, both in terms of commuting cost as seen above as well as in terms of evading tax and contributions over public utilities such as lighting, water and the property tax. Therefore, it is unclear if the increase on the value of the properties would really increase the wealth and income of current residents of if it would only compensate for the increase in the cost-of-living driven up for the recovery cost of upgrading programmes.

1.26 Other outcomes that can be affected by the infrastructure upgrading of the *favelas* are positive externalities in relation to health. It is well-known the link between sanitation and provision of quality of water (as well as quantity) in preventing a series of diseases, particular, vector-borne disease. Additionally, since these diseases are one of the main causes of child mortality, it would be interesting to investigate whether the urbanization programme had positive effects on reducing child mortality.

II. FAVELA-BAIRRO

2.1 The *Favela-Bairro* program was an initiative of the Municipio do Rio de Janeiro. It grew out of the city's master urbanization plan of 1992, the *Plano Diretor*. The program was financed partly with IDB funds, in the form of two investment loans approved in 1995 and 2000 and partly with the municipality funds, each for 180 million dollars (with 120 million in counterpart). It was coordinated by the newly created *Secretaria Municipal de Habitação* (SMH), but involved other municipal agencies including the labor and social development secretariats, the municipal rubbish collection company (COMLURB) as well as agencies at the state level—most notably the state water company, CEDAE.

2.2 The scope of *Favela-Bairro* was medium-sized communities. As was mentioned in the introduction, the Plano Diretor of Rio contemplated a number of different urbanization programs. *Favela-Bairro* was designed to address the needs of communities with between 500 and 2,500 households. The city also implemented programs to deal with smaller favelas (*Bairrinho*) as well as programs to urbanize the city's largest *favela* communities (*Grandes Favelas*). In terms of the amounts per beneficiary, the *Favela-Bairro* program was the most ambitious, with an estimated expenditure per family of 3,500 dollars[26].

2.3 The objective of the two programs was to improve the living conditions of the urban poor, and in doing so the program contained a rather wide mix of different social infrastructure, land tenure, and social development components. Specific objectives included the reduction in the risk of geological and environmental accidents (mostly landslides and floods), increased transit access, reduction in the incidence of vector-borne disease, as well as increases in utilization of public services. The basic infrastructure

[25] The *PCBR* reveals that 46% of the head of household say that they chose to leave in a *favela* due to housing costs. Another 28% say that they live there because it is the same place where friends and relatives live (possibly, a measure of social capital), only 11% point to proximity of workplace as the main reason to live in a *favela*.

[26] This amount was later increased further to 4,000. For details see the IDB, "Project Completion Report: Popular Housing Urbanization Project of Rio de Janeiro", December 2001.

component consisted of installing water, gutter, sewerage and lighting hardware, as well as road improvements. This infrastructure was made operative with municipal and state provision of these public goods, including rubbish collection. The social component varied over time, and originally included the construction of early child care centers. However, most social programs were delayed until the second phase and they were *de facto* absent from the first phase of *Favela-Bairro*, which is the focus of this evaluation. In the first phase, however, training programs were included with other funds, namely from the *Fundo de Amparo ao Trabalhador* (FAT). The project also involved the community in the selection of projects, although the exact mechanism by which these choices were made, as well as the *de facto* magnitude of local control, were not made clear during the evaluation interviews.

2.4 The selection of communities into the program and the order in which *favelas* were chosen in the project is particularly important, since this is the key to any evaluation that deals with endogenous program placement. The loan document included provisions that objective criteria be developed for the selection and ordering of communities in the project—and this was done during project preparation and the results of the ordering were presented in the form of a community matrix[27]. These criteria were broadly based on the socio-economic deficit of communities, the subjective assessment of the difficulties of implementing public works, and the public works (water, sewer, rubbish collection) deficits. In all 89 communities are identified in the loan document's matrix. However, the Municipality developed its own criteria for selection which do not match the IDB criteria. The data reviewed suggests that the IDB criteria was used to rank *favelas* conditional on program participation, while the Municipality criteria were used to select communities into the program. This point will be discussed in more detail in the section dealing with methodology.

2.5 In phase two of the project the social components were strengthened, and some that were originally programmed for the first phase were implemented. The program expanded so that child care centers were constructed, the training and community activities program was expanded, and specific modalities were added, including the creation of *agentes comunitarios*. These were members of the *favela* who were trained in matters of community development, hygiene[28], as well as in the specifics of the program[29]. Another difference between the two phases is that in the second phase more irregular settlements outside the *favelas* were scheduled to be urbanized, and a property titling program, which had been initially planned for the first phase, began execution in the second phase.

2.6 The changes from phase one to phase two, as expected, also come with changes in specific objectives. Whereas in the first phase the project was mostly focused on narrowly defined socio-economic outcomes in health and public services, in phase two

[27] In section IV we discuss the nature of selection, and how the rules that affected selection can be used to identify program participation.

[28] According to the Project Completion Report approximately 211,000 *favela* residents (in 38 *favelas*) and 51,000 residents (in 8 lotes) in other irregular settlements received project benefits. One must note that land titling has advanced fastest among the non-*favela* residents, since their land was purchased from legal plots of land that had been sub-divided illegally, as opposed to *favelas* which are mostly illegal occupations of private and public land.

[29] The *gari comunitario*, or community garbage collector, was also trained and hired from within the community.

project's objectives expanded appreciably, as did the set of indicators to be measured. In phase two indicators relative to child and youth education and health, housing value, and integration with the city were introduced[30].

2.7 In terms of outputs, the first phase of *Favela-Bairro* was able to accomplish most of what it set out to do. Two hundred and eighty-four public works and other projects, accounting for over 90 percent of the programmed activities, were executed in the target *favelas*. In all, 38 of the targeted 54 *favelas* were intervened. The program was also responsible for including the improved communities into the city map, and in the process providing basic sanitation services. Of the major components only the monitoring and evaluation components were not executed: the programmed evaluation activities were delayed, so that information immediately after the completion of works was not collected[31].

2.8 The difficulties with the data notwithstanding, the evaluation strategy of both the first and the second phase of the program did not explicitly provide for a comparator group that would form the basis for a counterfactual in any impact evaluation. In all cases, the data collected were in communities that were receiving or were scheduled to receive project benefits. This included the data generated in the PCBR, the data collected by DATABRASIL, as well as the data collected by AGRAR as part of the baseline for the second phase of the project. The consequence of this is that counterfactuals would have to be reconstructed *ex post*, based on a combination of household, census, and project-level data. These data issues, as well as the inescapable issue of endogenous program placement, are discussed in the next section.

III. EVALUATION STRATEGY: METHODOLOGY AND DATA

3.1 The definition of precise treatment and comparison group is one of the main problems regarding impact evaluation of social programs. This is so because one needs to find a counterfactual (given by a comparison group) that could ideally represent what the output (or outcome) would have been for the treated group in the absence of the program. Without this counterfactual very little can be said about the impact of the program, always taking impact to mean the causal effect of the program. Several techniques have been developed to tackle this issue, especially when good data is not available. In general, *ex post* evaluations have to use all sources of data available (surveys, administrative data, census) in order to try to build a credible counterfactual in the

[30] Indicators were evenly divided between opinion poll data and objective measures.

[31] There was an initiative to collect data, hired by the SMH through the firm DATABRASIL. However, due to problems between the SMH and DATABRASIL this data was eventually lost; it had to be recovered as part of this evaluation exercise.

absence of program specific surveys with both baseline and follow up interviews for treated and control groups alike.

3.2 The most credible evaluations try to build comparison groups that most resemble the treated group in both observable and unobservable dimensions. In circumstances where data is available before and after the intervention, and for which the treatment group represents a subset of the population, propensity score matching with differences-in-differences is typically used to model the impact of the treatment on the treated. Conditioning on the propensity score can potentially make the comparison group as similar as possible to the treated group as far as observables are concerned. As for the unobservable dimensions, the bias caused by the correlation between program assignment and the outcome (or output) can be wiped away via differences-in-differences, assuming that unobservables correlated with outcomes are time invariant.

3.3 The logical framework of *Favela-Bairro*[32] allows us to identify two outcomes and three outputs. The outcomes are (i) reduction of the incidence of diseases caused by lack of sanitation, (ii) reduction of occurrences of land slide; and the outputs are (i) to improve water access, connection to sewer, access to rubbish collection and power supply, (ii) to facilitate the access of the benefited population to other areas of the communities and to the rest of the town through paved streets and way-out; (iii) to increase the supply of nurseries. The general goal was to improve the living standards of the low-income population of Rio de Janeiro. In addition to the outcomes listed above we can also reasonably expect that the project would have a number of unanticipated results. Infrastructure improvement could have an impact on the value of properties, to the extent that housing value is determined by amenities and public services[33]. Education and income-related outcomes could be also be associated with the program, particularly since initiative in these sectors were incorporated into *Favela-Bairro* with funds from outside the program. In particular, several training courses financed by the *FAT* were developed in intervened communities[34].

3.4 The literature on program evaluation points to experimental designs as the most effective manner to carry out an impact evaluation. This occurs because the randomization of the treated group and of the control group implies that a counterfactual can be built based on the control group so that both observable and non-observed characteristics are balanced between the two groups. Since the control group is in every aspect similar to the treated one, every difference in the evolution of the evaluated outcome and outputs after the program can be attributed to the program (treatment).[35]

3.5 However, experimental design is not always feasible, due to cost considerations, the political and ethical difficulties in leaving out potential beneficiaries of the program, or,

[32] See IDB, "Brazil: Rio de Janeiro Urban Upgrading Program (BR-0182) Loan Proposal", 1995.
[33] These indicators were introduced in phase II.
[34] Those initiatives were later included in the design of PROAP-II.
[35] This result depends on two assumptions: (i) that the occurrence of drop-outs follows a random process and (ii) that members of the control groups do not attend (or be subject to) a program similar to the one they were excluded. See Manski (1996). This latter assumption is particularly important in the case of PROAP I evaluation since some control neighborhoods may have received a similar treatment. We took care of it by excluding communities that received programs such as *bairrinho* and *grandes favelas* which had similar characteristics to PROAP I.

perhaps what is most common, lack of interest or technical expertise by the part of policy-makers designing the project. In this circumstance the literature acknowledges that the second best option is a quasi-experimental design where we can build *ex post* a comparison group that most resembles the treated one in several dimensions.

3.6 We can build several comparison groups in order to estimate the average effect of the treatment on the treated population: $E(Y_{it}^1 - Y_{it}^0 \mid I = 1)$, where Y is the outcome (or output) variable and the superscript indicates the participations on the program, 1 for participants and 0 for non-participants. The missing observation on the expression above is $E(Y_{it}^0 \mid I = 1)$, i.e., the average of the outcome had the treated group *(I=1)* not participated on the program. If we use as a counterfactual – to replace the missing data – only the information on the outcome before the program for the participants, $E(Y_{it-k}^0 \mid I = 1)$, where t<k, one could argue that the estimated average difference would occur even in the absence of the program, and, therefore, cannot be attributed to it. The technique of reflexive comparison potentially leads to upward biased estimates because it cannot separate out what is the effect of the program (intervention) and what is the effect of macro changes or trends[36]. Therefore, the implementation of surveys that collect data before and after the intervention for treated communities such as the ones collected by DATABRASIL and AGRAR to evaluate PROAP I is not enough to establish causal links between the program and changes in the output and outcomes. In order to establish that causal link, data on both treated and non-treated communities before and after the treatment is a necessary condition to be sure that the typical bias of a reflexive comparison will not contaminate the results. In order to control for this bias, the difference of the before-after differences for the two groups yields the best possible estimator (assuming that a good matching was possible between treated and comparison groups) of the impact of the program in the absence of an experimental design. This is so because we can control both the selection into the program due to non-observable and time invariant factors that are wiped out in the first difference (reflexive comparison), and also the observed characteristics (through the matching). The double difference also cancels out the bias due to changes in economic environment, i.e., common shocks that may have affected outcomes as well as the selection variables of both treated and comparison groups and that are not dealt with in a simple before and after (analysis)[37]. Therefore, the difference-in-difference estimator with a comparison group defined via propensity score can be written as:

$$E(Y_{it}^1 - Y_{it-k}^1 \mid P(X), I = 1) - E(Y_{it}^0 - Y_{it-k}^0 \mid P(X), I = 0) \qquad (3.1)$$

where *P(X)* is the propensity score estimate.

3.7 This is not the first impact evaluation of PROAP-I based on census and with a difference-in-difference approach. IPP (2003)[38] presented a before-and-after evaluation for both

[36] Under the pessimistic assumption that things are progressing negatively over time this bias may in fact be negative (underestimated).

[37] This result hinges on the assumption that both treated and non-treated group had a similar trend before the treatment.

[38] See Instituto Pereira Passos, "O Momento 2000 do *Favela-Bairro*: Avaliação com Base nos Censos 1991 e 2000", 2003.

treated and comparison groups. However, the number of comparison group was limited to 17 based on a group of communities that were selected to take part in the second phase of the project (PROAP II). Thirty-seven treated communities were considered in this study out of the 54 communities that took part in the first phase of the program. This reduction was necessary due to the absence[39] of some treated communities in the Census Data (such as *Vila Campinho* and *Morro do Escondidinho*) and due to the fact that some of the improvements had not been implemented by July of 2000, when the census interviewers started going to the field. The comparison group was deemed to be a good one for it was selected using the same criteria of the communities that took part in the first phase. The output variables analyzed in that study were the ones related to the infrastructure improvements: water access, access to sewer, and rubbish collection. The main findings of the IPP study were that changes in the head of household education, changes in illiteracy rate, and changes in the income of the head of the household were the same in program and non-program communities. Changes in outputs such as water, sewer and rubbish collection were greater in program than in non-program communities. The IPP study, however, did not take into account possible sources of error (other than sampling error, since in a census there is none), and did not provide standard errors for the estimated mean differences.

3.8 Our study aims to tackle two aspects that were not dealt with by IPP's work: (i) to define comparison group based on a statistical procedure that allow us to match treated and comparison groups in the best possible way[40] and (ii) to test the statistical significance of the differences between means of treated and comparison communities. To do this we first define the comparison group based on a propensity score analysis. We will use four different estimates to measure the impact of the treatment in a difference-in-difference framework, being three of them based on a propensity score analysis. In the Nearest Neighbor (NN) version for instance each treated community is matched to a comparison community with the closest propensity score. For each period, before and after the treatment, the difference between treated and comparison groups can be written as: $a_{mcs} = \frac{1}{n_1}\sum_{i \in I=1}[Y_i - \sum_{j \in I=0} W_{ij} Y_j]$ and their difference may be calculated as $a_t - a_{t-k}$[41], where *mcs=t or t-k* indexes the matched cross section, either before *(t-k)* of after the treatment *(t)* and W_{ij} corresponds to the weight used for the comparison group *(j)*. In the case of the nearest neighbour W_{ij} is equal to one, whereas for extended versions of NN, say T-NN, W_{ij} = 1/T. In the empirical part of this paper we will use the nearest neighbour standard version and the 5-NN. The third alternative is the Kernel matching. The kernel matching builds the counterfactual based on a kernel weighted average over the set of communities in the comparison group:

$$a_{km} = \frac{1}{n_1}\sum_{i \in I=1} Y_i - \frac{\sum_{j \in I=0} Y_j K(\frac{P_j - P_i}{h_n})}{\sum_{k \in I=0} K(\frac{P_k - P_i}{h_n})} \qquad (3.2)$$

[39] Actually those communities were not identified in the file with the name of the slums yielded by the statistics office with the documentation for the 1996 Population counting.

[40] This is especially important because as we will show in the next section assignment rules were not very clear.

[41] Note that this difference is equal to the one in (3.1).

where $K(.)$ is a kernel function and h_n is a bandwidth parameter and P_i is the propensity score for individual i. The weight, W_{ij}, corresponds to $\frac{K(\frac{P_j - P_i}{h_n})}{\sum_{k \in I=0} K(\frac{P_k - P_i}{h_n})}$ in the kernel matching. Heckman et al. (1997) show that under standard conditions on the bandwidth and the kernel[42], the counterfactual generated by this method is a consistent estimator of $E(Y^0 \mid P(X), I = 1)$. The choice of the kernel and of its bandwidth will determine the number of comparison observations used to match each treated individual and the weight[43] with which each control observation will enter the calculation of the counterfactual. The Gaussian kernel, for instance, is unbounded which means that all control observations enter the calculation. In this paper we use the Epanechnikov kernel, which is bounded, with constant bandwidth of 0.06.

3.9 The data to be used in this evaluation comes from several sources. The output and main socio-economical features of treated and comparison communities come from two main sources: the Population Census and the more detailed 10% sample of the Population Census. The questionnaires of the General Population Census and the 10% sample differ in relation to the number of variables investigated[44]. Unfortunately, the questionnaires are not entirely comparable between 1991 and 2000, despite investigating mostly the same variables[45]. Access to water, for instance, is captured in two questions in 1991 and only one in 2000. Nevertheless, it was possible to make the definitions compatible in the two censuses.

3.10 In order to get data for rents after 1996 – as we did not have this information on the 2000 Population Census - we had to rely on the data provided by the Survey on Low Income Community Survey (PCBR), and from data from the PNADs. The PNAD datasets were already described in section II. The PCBR was carried out by the Scientific Society of the National School of Statistics Science (SCIENCE), at the request of the Municipal Labor Office and was paid with funds from FAT. The purpose of the study was to identify the needs of the different communities that had been assigned to PROAP I. It was part of the strategy of the Municipality to take other social services, mainly, training to those communities in order to complement the infra-structure works of *Favela-Bairro*. The survey was carried out in 54 communities[46].

3.11 Rent information (and property information more broadly), was also available from the annual PNAD surveys. The drawback of the PNADs is that comparisons between *Favela-Bairro* communities and other *favelas* are not possible, since identification of

---

[42] These conditions require that $K(.)$ integrates to one, has mean zero and that $h_n \to 0$ as $n \to \infty$ and $nh_n \to \infty$.

[43] The weight of each comparison observation decreases with its distance from the treated observation as measured by their propensity scores.

[44] The latter brings information on durable goods, on migration pattern, religion, schooling and working status.

[45] The 10% sample of the 2000 Population Census does not investigate the value of rents paid and does not bring information on the quality of the roof and walls of households which could give us a hint of how improvements in treated communities differed from improvements in comparison ones between 1991 and 2000.

[46] They are listed in the appendix.

beneficiary communities is not feasible given the identifying information in the dataset. The only comparisons that can be made using the PNADs are comparisons between all *favelas* in the surveys and other low-income communities of similar characteristics. Also, given that we have data on all years of implementation up to 1999, we identify the control group through the probit equation estimated in Table 1.3, and select as a control group only those communities in the common support of this probit.

3.12 Death Certificates are another important source of information for this evaluation. This information is used mostly to assess the program's impact on different types of death causes, including those related to lack of sanitation and those related to homicide. The data was collected from the Municipal Health Department and it is judged to be of high quality. More details about the kind of information it provides will be presented on section about the evaluation of health impacts of the program.

3.13 Besides the data described above, data was also specifically collected for the evaluation of *Favela-Bairro*. This monitoring and evaluation system was part of the contract between the Mayors's Office and the IDB. The SMH contracted the firm DATABRASIL to carry out a field survey to capture characteristics of communities that were to be treated—this was dubbed the Moment 0 (M0). However, delay on the contracting process implied that some communities were interviewed when the works had already started. In addition to M0, the project provided for an assessment of beneficiaries in M1 and M2000, which should have consisted of evaluations based on a follow-up survey (M1) and on the census data (M2000) when the first results where released. M2000 was carried out by IPP and consisted on the evaluation that we discussed in the last section. The M1 survey was scheduled to go to field within six months of benefits. However, due to problems in selecting a firm to conduct the survey it only went to field in 2003 and the results are currently being tabulated by AGRAR, the firm that was hired to conduct it[47].

3.14 DATABRASIL questionnaires were not homogeneous in all surveyed communities. We managed to identify at least 3 types of questionnaires. Nevertheless, they follow a common pattern. They have similar question to the ones available in the Population Census, but more detailed information on property ownership and certificate, questions about community life and participation on the social movements, as well as, a subjective evaluation of the facilities and public services available in the community.

3.15 The goal of these surveys was to generate information on M1 as well as compare M0 to M1. For this latter goal the AGRAR questionnaire was designed to include some of the same questions that were fielded by DATABRASIL. The problem here is that only 5 communities (one of them being a complex of 4 communities) were surveyed in both M0 and in M1. There was also no attempt to sample from non-treated communities, so that the only types of comparisons possible with these data would be a reflexive comparison of a reduced subset of the communities intervened.

[47] As mentioned in section three, the microdata of DATABRASIL (M0) was lost by the SMH and at the moment efforts are underway to recover this data. The files that were made available so far are not documented in a manner that would make it comparable to the ones provided by AGRAR (M1).

IV. RESULTS

4.1 This chapter presents the main results of the impact evaluation. Further detail can be found in the appendices. These consist of both results using propensity scores (described above) as well as reflexive before and after comparisons. In all cases, the bias associated with different evaluation strategies will be related to the fundamental problem of endogenous program placement. In other words, if communities are selected for treatment non-randomly—as may be the case here if one takes the *favela* matrix described in section III—the results after program implementation may be as much an artefact of this non-random selection that of the program's true impact.

4.2 In order to properly account for the nature of program placement one must understand the process in which communities were selected into *Favela-Bairro*. At least initially this process seems to be clear cut, since it is spelled out in the Loan Document prepared by the IDB: communities are selected in a first stage based on their size (only those between 500 and 2,500 are eligible) and based on the cost of urbanization (those with extremely high costs are excluded), and in a second stage based on infrastructure deficits, social deficits and the ease of completing infrastructure works[48]. However, the criteria documented by the SMH are somewhat different, since they do not include, for example, the social deficit score. In addition, it is not clear if even the sign with which the criteria are used to rank the communities was the same in both methods. A review of the available documentation, however, suggests that one set of criteria was used in the selection of intervention communities and a second set was used in ranking them in the *Favela-Bairro* matrix[49]. In any case, the variety of potential methodologies used in selection and in ordering motivated an *ex post* assessment of the characteristics of both program and non-program communities, i.e. a modelling of program participation.

4.3 Program participation was modelled using a probit specification in order to understand the influence of each criterion on the selection procedure. Table 1 presents the marginal effects of these probit models. Column [1] shows the coefficients for the criteria indicated by the IDB. Only the percentage of illiterate head of the household had a statistically significant impact on the probability of being selected to take part in the program, all the other variables were not statistically significant and some of them even show a wrong sign. The Likelihood ratio test for the joint significant of the model also shows that jointly those variables are not statistically significant to explain selection into the

[48] According to the IDB documentation, the order of *favelas* would be based on the infrastructure deficit as measured by the arithmetic average of (i) % household without water access, (ii) % of household without proper sewage and the social and on the socio-economical deficit as measured by the arithmetic average of (i) % of children under 4 years old, (ii) % of female household heads, (ii) % household head with monthly income below the minimum wage, e (iv) % of illiterate household head.

[49] According to the methodology identified by the SMH the criteria for selection would be: water access, proper sewage, draining infrastructure, land-slide and flood risk, paved streets, and spill-over potential. All communities that were eligible for the program (i.e. that had between 500 and 2,500 households) were ranked in a 0 to 10 scale according to those six criteria. These ratings were done by a group of experts from several municipal offices. According to this methodology priority would be given to communities that would be easier to urbanize, either because of access or because of existing infrastructure. See the SMH document "Documento 2: Metodologia de Classificação de *Favelas* (proposta preliminar)", no date.

program. Columns [2], [3], and [4] reveal that the most important factor to explain selection into the program is actually the size of the community. Given the threshold of 500 households and the ceiling of 2,500, it does not come as a surprise that the size of the community and its square (this one with a negative sign) are both highly significant to explain the selection. The models that do not include the size of the community are never statistically significant regardless of the set of variables used to explain the selection, see Columns [1] and [5]. If one looks at the household level of each community in order to predict participation results are similar. However, in this case we see some evidence that the IDB criteria were taken into account, but the impact is very small (see Table 4.2). It is worth noting that conditional on program selection, the communities with the largest socio-demographic deficits were, on average, slated to be treated before the other communities. In other words, although the program did not target the neediest communities for possible intervention, they did give more needy communities priorities in the order of execution of public works, conditional on inclusion in the intervention matrix.

4.4 From an evaluation point of view the fact that community selection is mostly uncorrelated with the purported selection criteria is actually a positive finding, to the extent that it may attenuate the bias associated with selection. This is, of course, under the hypothesis that selection was not over unobservables that may, nevertheless, be correlated with outcomes. This hypothesis is always a potential problem with non-experimental design, and ruling it out entirely is conceptually impossible.

4.5 From the aggregate results we can reasonably assume that the selection process conditional on the size of the community was either almost random or that non-observed factors played a major role. Given this last possibility and the different results that we get when household microdata is used instead of community level ones, we will use the largest set of information available on the Population Census both regarding both infrastructure deficit and social and economical deficits of the eligible communities in order to build the best comparison groups via propensity score matching. We are aware of the problems involved in using output measures into the matching criteria, but since they were stated as ranking criteria, we decided to use them[50].

[50] Results were the matching process is done without using output variables will also be reported for the sake of comparison.

Table 4.1 - Evaluating Community Selection Criteria (Marginal Effects of Probit Model)

	[1]	[2]	[3]	[4]	[5]
% water network access	-0.068	-0.054		-0.05	-0.084
	[0.058]	[0.034]		[0.031]	[0.060]
% sewer network access	0.026	0.005		0.006	0.01
	[0.039]	[0.020]		[0.018]	[0.040]
% children younger than 4 years old	-0.303	0.122	0.223	0.405	0.553
	[0.651]	[0.404]	[0.615]	[0.643]	[1.064]
% female head of household	0.219	0.22	0.143	0.106	0.02
	[0.206]	[0.134]	[0.136]	[0.133]	[0.240]
% head of household with income lower than 1 mw	-0.079	-0.052		-0.107	-0.222
	[0.092]	[0.047]		[0.080]	[0.159]
% iliterate head of the household	0.653	0.348		0.012	0.491
	[0.208]**	[0.141]*		[0.313]	[0.557]
ln (size of the community)		0.387	0.352	0.334	
		[0.083]**	[0.084]**	[0.084]**	
ln (size of community) squared		-0.029	-0.026	-0.025	
		[0.007]**	[0.007]**	[0.007]**	
% head of household with less than 4 years of schooling			-0.066	0.004	-0.097
			[0.191]	[0.182]	[0.380]
% illiterates among 15 years old and more			0.31	0.276	0.188
			[0.248]	[0.388]	[0.688]
residents per household			0.061	0.061	0.051
			[0.031]	[0.030]*	[0.051]
% female population			0.642	0.588	1.433
			[0.533]	[0.507]	[0.801]
averare years of schoold of the head of the household			-0.026	-0.013	-0.026
			[0.029]	[0.029]	[0.055]
average age			0.003	0.006	0.013
			[0.007]	[0.007]	[0.012]
own house			-0.07	-0.063	-0.088
			[0.061]	[0.056]	[0.110]
ln (real average income of the head of the household			0.029	-0.042	-0.108
			[0.031]	[0.054]	[0.111]
% rubbish collection				0.017	0.08
				[0.030]	[0.063]
Observations	435	435	435	435	435
Chi2	12.34	73.44	77.15	82.8	20.76
Pseudo R2	0.04	0.24	0.25	0.27	0.07

Standard errors in brackets
* significant at 5%; ** significant at 1%

[1] Corresponds to a model that takes into account only IDB criteria
[2] Includes the size of the household into [1]
[3] Excludes ouput variables and Includes more socio-economical variables
[4] Equal to [3] including output variables
[5] Equal to [4] but excludes the size of the household

Table 4.2 - Evaluating Community Selection Criteria (Marginal Effects of Probit Model)

	[1]	[2]	[3]	[4]	[5]
% water network access	-0.122	-0.06		-0.059	-0.118
	[0.003]**	[0.002]**		[0.003]**	[0.004]**
% sewer network access	-0.009	-0.002		0	0.001
	[0.002]**	[0.001]**		[0.001]	[0.002]
% children younger than 4 years old	0.004	0.003	0.002	0.001	0.004
	[0.006]	[0.002]	[0.003]	[0.003]	[0.008]
% female headship	0.018	0.006	0.006	0.006	0.018
	[0.002]**	[0.001]**	[0.001]**	[0.001]**	[0.003]**
% head of household with income lower than 1 mw	-0.021	-0.009		-0.008	-0.027
	[0.002]**	[0.001]**		[0.001]**	[0.003]**
% iliterate head of the household	0.003	0.003		0.004	0.011
	[0.002]	[0.001]**		[0.002]**	[0.004]**
ln (size of the community)		1.071	1.065	1.049	
		[0.013]**	[0.015]**	[0.015]**	
ln (size of community) squared		-0.067	-0.067	-0.066	
		[0.001]**	[0.001]**	[0.001]**	
% head of household with less than 4 years of schooling			0.004	0.004	0.008
			[0.002]*	[0.002]**	[0.004]*
% illiterates among 15 years old and more			0.006	0	-0.004
			[0.002]**	[0.002]	[0.005]
residents per household			0.002	0.002	0.009
			[0.000]**	[0.000]**	[0.001]**
% female population			0.001	0.002	0.008
			[0.002]	[0.002]	[0.004]
averare years of schoold of the head of the household			0	0.001	0.002
			[0.000]	[0.000]**	[0.001]**
average age			0	0	0
			[0.000]	[0.000]*	[0.000]**
own house			-0.001	-0.001	0.021
			[0.001]	[0.001]	[0.003]**
ln (real average income of the head of the household)			0.005	0.002	-0.002
			[0.001]**	[0.001]*	[0.002]
% rubbish collection				-0.007	-0.037
				[0.001]**	[0.003]**
Observations	223559	223559	150319	150319	150319
Chi2	2300.82	12891.84	8311.61	8778.37	1850.73
Pseudo R2	0.01	0.22	0.21	0.22	0.01

Robust standard errors in brackets
* significant at 5%; ** significant at 1%

A. Evaluating Outputs and Outcomes Using the Population Census

4.6 The propensity score is constructed using a logit model with data from the 1991 census as the baseline of this study[51]. In addition to the analysis using the census track level census data, we also estimate models using the microdata from the 10 percent sample of the census. When analysing the more detailed information of the 10% sample of the population census we are able to use individual level information and for this reason to analyse whether there has been any heterogeneity on the impact of the program in relation to the position of the beneficiaries in the income distribution of the treated incomes.

4.7 In order to evaluate output delivery and outcome impacts based on the population census, we selected three outputs: water access, sewer access and rubbish collection, and three outcomes: illiteracy among heads of household, income of the head of the household and increase in the population of the community.[52]

4.8 Following IPP documentation we will consider only 38 communities of the first phase for which the works had been completed by July 2000 and which were identifiable in both 1991 and 2000 population censuses. We also excluded communities that took part in another urbanization program (*bairrinho*) and the ones who were considered not suitable for such program due to its proximity to risky areas (e.g. railway margins) or to its environmental potential damage (e.g. too close to water supply sources).

4.9 The regressors used in calculating the propensity scores were size of the slum, water access, sewer connected to the city, rubbish collection, proportion of people with less than 4 years of schooling, proportion of people with more than 15 years of schooling, overall illiteracy rate, head of household illiteracy rate, proportion of people with less that 4 years of age, proportion of women headship, proportion of head of household earning less than 1 minimum wage, proportion of women, average years of schooling, average age, ownership of the house, average income, dummies for administrative areas. We run four specifications in order to assess the robustness of the results. The main variant in calculating the propensity score is the inclusion or exclusion of output and outcome variables in the logit specification. Including the outputs and outcomes can be justified on the grounds that those variables may have been factors in the selection of communities and in the order of their incorporation into the program. In principle the inclusion of outputs and outcomes can be problematic given that it may lead to a violation of a main assumption required in order to identify the average treatment effect on treated in the literature on program evaluation: namely the conditional mean independence assumption, which maintains that conditional in a rich set of covariates, assignment to the program is

[51] Unfortunately, the 1996 Population counting did not collect detailed information on household infrastructure and on several important individual characteristics. For this reason we have to rely on a baseline that is almost 5 years old in relation to the date of the start of the program.

[52] Actually those outcome are not closely linked to the one established on the aims of the program. Nevertheless based in our discussion in last section we could reasonably assume that the program may have had an impact on both literacy and income via the complementary programs put in place by the municipality on those treated communities. Moreover, the existence of the program may have had some migration effect.

mean independent of its outcomes in *t* and *t-k*[53]. For this reason we also estimate propensity score without including those output and outcome variables. We also run a full sample model as well as a model with only eligible communities, for robustness checks. As our baseline is set well before the program starts we decided to be more flexible in relation to the cut off points and included communities between 100 and 2600 households, instead of sticking to the 500 and 2500 interval.[54]

4.10 In analyzing the results we have to keep in mind the small size of treated observations (only 38 communities) against a total of comparison observations that vary from 38 in the case of the Nearest Neighbor procedure to 272 in the case of the regression specification. Such small number of observations may produce imprecise estimates; therefore the sign of the estimate should be considered even when standard errors are reasonable, yet not small enough to allow for significance at conventional levels. In the case of the propensity score matching estimates, the standard errors were calculated via bootstrap with 1000 replications. To account for the problems of the small number of observations, mainly in the case of the Nearest Neighbor procedure, we report the standard errors with strata option, in order to guarantee that 38 treated observations will always be present in the replications. Table 4.3-4.5 present the results for all specifications, in addition to the difference-in-difference results (diff-in-diff), the specification based on propensity score matching (Tables 4.3 and 4.4) also report the average treatment effect based on a simple cross-section for 2000 (ATT 2000) and the difference before the program (ATT 1991).

4.11 As for the specifications based on propensity score matching that use output and outcomes in the calculation of the propensity score, Table 4.3 shows that only the difference-in-difference estimates for sewer access achieve statistical significance. The difference-in-difference estimates show a positive effect that varies from 17% to 23.5% in terms of increase in access to the general sewage network in treated communities. The effect on the other two outputs seems to be positive but not statistically significant for water and very imprecise for rubbish collection. The estimates for water access vary from 6.4% to 8.1%. As for rubbish collection, the estimates vary from –2.9% to 7.3%. Similarly, the estimated effect on the three outcomes—illiteracy rate of head of household, income of the head of the household and population—is never significant. However, whereas, the impact on illiteracy shows in general a positive effect, the impact on income and on the population size is mostly negative, regardless of using the full sample or the sample of eligible communities by size.

[53] See Rosenbaum and Rubin (1983).

[54] It is worth mentioning that there was some exceptions and very small communities were also treated, mainly when they conurbate with other communities.

Table 4.3 - Impact On Outputs And Outcomes
(Specification using outputs and outcomes on the propensity score)

FULL SAMPLE		water	sewer	rubbish	illiteracy	income	population
Nearest Neighbour	Diff-in-Diff	0.081 [0.087]	0.193 [0.128]	0.024 [0.068]	-0.002 [0.019]	-0.121 [0.093]	-0.164 [0.268]
	ATT 2000	0.103 [0.078]	0.291 [0.093]*	0.030 [0.034]	0.009 [0.018]	-0.113 [0.084]	0.908 [0.433]*
	ATT 1991	0.022 [0.086]	0.098 [0.109]	0.005 [0.078]	0.011 [0.021]	0.008 [0.090]	1.072 [0.443]*
5 - Nearest Neighbour	Diff-in-Diff	0.079 [0.057]	0.187 [0.084]*	0.042 [0.043]	0.013 [0.013]	-0.068 [0.067]	-0.161 [0.165]
	ATT 2000	0.045 [0.051]	0.196 [0.065]*	0.030 [0.023]	0.007 [0.013]	-0.050 [0.059]	0.830 [0.288]*
	ATT 1991	-0.034 [0.050]	0.009 [0.069]	-0.012 [0.045]	-0.006 [0.013]	0.018 [0.054]	0.991 [0.277]*
Kernel	Diff-in-Diff	0.064 [0.057]	0.170 [0.083]*	0.051 [0.044]	0.014 [0.013]	-0.080 [0.070]	-0.203 [0.174]
	ATT 2000	0.040 [0.052]	0.170 [0.061]*	0.024 [0.021]	0.008 [0.013]	-0.052 [0.060]	0.774 [0.288]*
	ATT 1991	-0.024 [0.051]	-0.001 [0.069]	-0.027 [0.048]	-0.006 [0.012]	0.028 [0.058]	0.977 [0.278]*
Nearest Neighbour -	Diff-in-Diff	0.081 [0.079]	0.193 [0.129]	0.024 [0.066]	-0.002 [0.019]	-0.121 [0.094]	-0.164 [0.275]
(bootstrap with strata)	ATT 2000	0.103 [0.079]	0.291 [0.094]*	0.030 [0.033]	0.009 [0.018]	-0.113 [0.082]	0.908 [0.417]*
	ATT 1991	0.022 [0.084]	0.098 [0.112]	0.005 [0.074]	0.011 [0.021]	0.008 [0.093]	1.072 [0.424]*
ELIGIBLE SIZE		water	sewer	rubbish	illiteracy	income	population
Nearest Neighbour	Diff-in-Diff	0.073 [0.096]	0.214 [0.155]	-0.029 [0.086]	0.007 [0.020]	-0.105 [0.107]	-0.155 [0.195]
	ATT 2000	0.106 [0.090]	0.233 [0.126]**	0.022 [0.039]	0.003 [0.024]	-0.063 [0.102]	0.117 [0.429]
	ATT 1991	0.033 [0.115]	0.019 [0.156]	0.050 [0.093]	-0.005 [0.027]	0.042 [0.121]	0.272 [0.403]
5 - Nearest Neighbour	Diff-in-Diff	0.065 [0.066]	0.220 [0.098]*	0.073 [0.051]	0.004 [0.012]	-0.061 [0.070]	-0.092 [0.141]
	ATT 2000	0.054 [0.057]	0.198 [0.079]*	0.024 [0.029]	0.002 [0.013]	-0.047 [0.060]	0.101 [0.261]
	ATT 1991	-0.011 [0.063]	-0.023 [0.084]	-0.049 [0.052]	-0.002 [0.013]	0.014 [0.065]	0.193 [0.211]
Kernel	Diff-in-Diff	0.075 [0.072]	0.235 [0.109]*	0.071 [0.060]	0.001 [0.014]	-0.045 [0.077]	-0.035 [0.142]
	ATT 2000	0.045 [0.063]	0.232 [0.087]*	0.021 [0.028]	-0.008 [0.016]	-0.033 [0.064]	0.238 [0.280]
	ATT 1991	-0.031 [0.071]	-0.002 [0.097]	-0.050 [0.061]	-0.008 [0.016]	0.012 [0.070]	0.274 [0.248]
Nearest Neighbour -	Diff-in-Diff	0.073 [0.101]	0.214 [0.151]**	-0.029 [0.077]	0.007 [0.020]	-0.105 [0.104]	-0.155 [0.204]
(bootstrap with strata)	ATT 2000	0.106 [0.090]	0.233 [0.126]**	0.022 [0.049]	0.003 [0.024]	-0.063 [0.093]	0.117 [0.443]
	ATT 1991	0.033 [0.110]	0.019 [0.157]	0.050 [0.083]	-0.005 [0.026]	0.042 [0.113]	0.272 [0.405]

Standard Errors into brackets

* Significant at 5% **Significant at 10%

Table 4.4 - Impact On Outputs And Outcomes
(Specification WITHOUT outputs and outcomes on the propensity score)

FULL SAMPLE		water	sewer	rubbish	illiteracy	income	population
Nearest Neighbour	Diff-in-Diff	0.058 [0.077]	0.235 [0.112]*	-0.008 [0.081]	-0.008 [0.018]	-0.226 [0.088]*	-0.053 [0.214]
	ATT 2000	0.025 [0.067]	0.070 [0.081]	0.020 [0.036]	0.001 [0.016]	-0.050 [0.079]	0.684 [0.421]**
	ATT 1991	-0.033 [0.073]	-0.166 [0.119]	0.028 [0.085]	0.009 [0.018]	0.176 [0.081]*	0.737 [0.393]**
5 - Nearest Neighbour	Diff-in-Diff	0.098 [0.064]	0.215 [0.086]*	-0.016 [0.061]	0.011 [0.013]	-0.130 [0.065]*	-0.092 [0.138]
	ATT 2000	0.047 [0.047]	0.173 [0.059]*	0.032 [0.024]	0.006 [0.011]	-0.065 [0.060]	0.809 [0.282]**
	ATT 1991	-0.052 [0.060]	-0.042 [0.089]	0.048 [0.065]	-0.004 [0.012]	0.065 [0.055]	0.901 [0.266]*
Kernel	Diff-in-Diff	0.105 [0.063]**	0.231 [0.084]*	0.027 [0.063]	0.006 [0.011]	-0.106 [0.062]**	-0.145 [0.135]
	ATT 2000	0.059 [0.047]	0.167 [0.051]*	0.030 [0.021]	0.011 [0.011]	-0.078 [0.052]	0.744 [0.278]*
	ATT 1991	-0.046 [0.058]	-0.065 [0.085]	0.002 [0.066]	0.004 [0.010]	0.028 [0.052]	0.889 [0.257]*
Nearest Neighbour -	Diff-in-Diff	0.058 [0.075]	0.235 [0.115]*	-0.008 [0.085]	-0.008 [0.019]	-0.226 [0.092]*	-0.053 [0.213]
(bootstrap with strata)	ATT 2000	0.025 [0.061]	0.070 [0.081]	0.020 [0.033]	0.001 [0.016]	-0.050 [0.075]	0.684 [0.404]**
	ATT 1991	-0.033 [0.074]	-0.166 [0.118]	0.028 [0.090]	0.009 [0.019]	0.176 [0.082]*	0.737 [0.386]**
ELIGIBLE SIZE		water	sewer	rubbish	illiteracy	income	population
Nearest Neighbour	Diff-in-Diff	0.125 [0.072]**	0.329 [0.111]*	0.011 [0.090]	0.004 [0.019]	-0.088 [0.010]	0.062 [0.175]
	ATT 2000	0.023 [0.057]	0.136 [0.094]	-0.002 [0.046]	-0.007 [0.016]	-0.041 [0.089]	0.253 [0.376]
	ATT 1991	-0.102 [0.072]	-0.193 [0.123]	-0.013 [0.094]	-0.012 [0.019]	0.047 [0.082]	0.191 [0.324]
5 - Nearest Neighbour	Diff-in-Diff	0.108 [0.064]**	0.333 [0.092]*	0.034 [0.071]	0.005 [0.012]	-0.076 [0.067]	0.037 [0.125]
	ATT 2000	0.013 [0.044]	0.167 [0.067]*	0.032 [0.034]	-0.001 [0.011]	-0.050 [0.058]	0.272 [0.232]
	ATT 1991	-0.095 [0.064]	-0.167 [0.097]*	-0.002 [0.072]	-0.006 [0.012]	0.026 [0.056]	0.235 [0.182]
Kernel	Diff-in-Diff	0.119 [0.069]**	0.335 [0.096]*	0.044 [0.071]	0.001 [0.013]	-0.070 [0.072]	0.018 [0.127]
	ATT 2000	0.020 [0.049]	0.182 [0.069]*	0.023 [0.040]	-0.003 [0.012]	-0.052 [0.059]	0.179 [0.264]
	ATT 1991	-0.099 [0.065]	-0.153 [0.102]	-0.021 [0.074]	-0.004 [0.014]	0.017 [0.057]	0.160 [0.222]
Nearest Neighbour -	Diff-in-Diff	0.125 [0.076]**	0.329 [0.115]*	0.011 [0.088]	0.004 [0.018]	-0.088 [0.096]	0.062 [0.177]
(bootstrap with strata)	ATT 2000	0.023 [0.056]	0.136 [0.093]	-0.002 [0.050]	-0.007 [0.016]	-0.041 [0.086]	0.253 [0.365]
	ATT 1991	-0.102 [0.073]	-0.193 [0.121]	-0.013 [0.092]	-0.012 [0.019]	0.047 [0.080]	0.191 [0.302]

Standard Errors into brackets

* Significant at 5% **Significant at 10%

4.12 The difference-in-differences results for the specifications that exclude outputs and outcomes from the propensity score are in general much larger, basically due to the lower quality of the matching, and for this reason a larger number of estimates are statistically significant. The estimates for the average treatment effect (ATT) in 1991, i.e., for our baseline, yield negative coefficients of greater magnitude in this specification than the specification that includes outputs and outcomes. However, they are only marginally statistically significant, with the exception of the estimates for population. The difference-in-differences estimates according to this specification show that both water and sewerage access to the general network increased substantially for treated communities vis-à-vis comparison ones. Estimates suggest that water access increased between 5.8% to 12.5% and Sewerage access increased between 21% and 33.5%. No effects are found for rubbish collection, illiteracy and population growth. In terms of income, for the full sample, the estimates suggest a significant negative, but we interpret these with caution given that they are not robust across empirical specifications, and in particular are sensitive to the variables used in the logit matching. Another empirical regularity of note in this specification is that it seems that there were instances of the "vanishing benefits" effect if we analyze only the average treatment effect after the treatment (ATT 2000), i.e. if we do not double difference the treatment effects. Several specifications of the cross-section matching for water and sewer access are positive but never statistically significant. However, the difference-in-difference estimates are positive and statistically significant.

4.13 Impact models were also estimated parametrically, primarily as a robustness check. The results in Table 4.5 are based on a regression model with interaction between the year 2000 and the treated communities in order to estimate the effects of the program. The parametric results do no differ from the ones obtained from the propensity score matching with difference in differences. There were effects on water and sewer access.

Table 4.5 - Regression Specification for Difference-in-Difference estimates

	Others Outputs and Outomes used as Full		Others Outputs and Outomes used as Eligble Size	
Water	0.066	[0.045]	0.067	[0.048]
Sewer	0.192	[0.073]*	0.189	[0.072]*
Rubbish	-0.026	[0.047]	-0.005	[0.044]
Illiteracy	0.000	[0.008]	0.006	[0.008]
Income	-0.012	[0.040]	0.000	[0.039]
Population	-0.238	[0.193]	-0.155	[0.106]
	No outputs and no Outomes used as Full		No outputs and no Outomes used as Eligble Size	
Water	0.085	[0.052]**	0.092	[0.049]**
Sewer	0.210	[0.075]*	0.220	[0.074]*
Rubbish	0.014	[0.049]	0.031	[0.045]
Illiteracy	0.001	[0.008]	0.005	[0.008]
Income	-0.018	[0.040]	-0.007	[0.039]
Population	-0.206	[0.192]	-0.124	[0.106]

Standard Errors into brackets

* Significant at 5% **Significant at 10%

The difference-in-difference estimates on water access varied from 6.6% to 9.2% and the one for sewer varied from 19% to 22% in line with the previous results.[55]

4.14 To illustrate the sort of bias we would get if we had implemented a reflexive comparison approach (before-after analysis), we report in Table 4.6 the results of a regression of outcome and outputs in relation to a year dummy (2000). The coefficient of the year dummy gives us the mean average effect of the program when we compare 2000 results with 1991 ones. The constant gives us the average level of outcome and outputs in 1991 for the 38 treated *favelas*. The results reveal that reflexive comparison would overestimate the actual effect of the program in more than 100% for water and sewer access and rubbish collection. The results for illiteracy and income of the head of household would be clearly biased upwards (in absolute value) and would be statistically significant. Clearly, Before-after analysis is a very poor indicator of the success of the program. These results based on census data make us very skeptical in relation to what kind of input longitudinal data only on treated *favela* can give us in an impact evaluation.

Table 4.6 Reflexive Comparisons Using Household Microdata

	Water	Sewer	Rubbish	Illiteracy	Income	Population
year==2000	0.139	0.436	0.188	-0.061	0.319	0.054
	[3.02]**	[6.96]**	[4.41]**	[5.37]**	[5.82]**	[0.34]
constant	0.805	0.442	0.793	0.202	5.733	7.823
	[24.74]**	[9.99]**	[26.37]**	[25.28]**	[147.82]**	[70.02]**
Observations	76	76	76	76	76	76
R-squared	0.11	0.4	0.21	0.28	0.31	0

Absolute value of t statistics in brackets

* significant at 5%; ** significant at 1%

B. Differences-in-Differences Using the Sample Census: Heterogeneity of Impact

4.15 The sample census can help us to check the results of the last section and also to evaluate if there was heterogeneity on the effect of the program. Basically we want to know if **there** was any difference in the effect of the problem according to the household income. Table 4.7 shows the results of the effect of the program on selected characteristics of treated communities. Due to the larger sample size - the unit of observation is not the community, but the individual – the coefficient of the interaction (and its correspondent marginal effect) are more robust than with aggregate data.

[55] It is worth noting that the specifications with outputs and outcomes as regressors obviously do not use the dependent variable, but use the other outputs and outcomes investigated on this paper.

Table 4.7 - Differences in Differences Estimates using household microdata

	water	sewer	rubbish	illiteracy	female headship	< 4 years of schooling
year (2000)	-0.009	0.130	0.109	0.039	0.030	-0.011
	[0.001]***	[0.002]***	[0.001]***	[0.002]***	[0.001]***	[0.001]***
treated	-0.087	-0.024	-0.021	0.006	0.001	-0.004
	[0.004]***	[0.004]***	[0.002]***	[0.005]	[0.003]	[0.003]
treated*year(2000)	0.056	0.158	0.040	-0.001	0.002	0.000
	[0.002]***	[0.003]***	[0.002]***	[0.006]	[0.004]	[0.004]

*** Signficat at 5%
sample weights used

4.16 The effects on water (6%) and sewer (16%) access, although marginally lower than the ones found in the previous specifications of the last section[56], point to a significant increase in water and sewer access in treated communities in comparison to non-treated ones. It is worth noting that for comparison communities the access to water was lower than the one observed in 1991 (the dummy for 2000 display a negative coefficient), in contrast sewer access also increased for them but at a lower rate than the one observed for treated communities. The only result that is slightly different from the ones reported in last section refers to the statistical significant increase in rubbish collection for treated communities. There was a 4% increase vis-à-vis the comparison group. The percentage of illiterates, the percentage of head of household with less than 4 years of schooling and female headship did not evolve in different ways in treated communities. In general, there was a significant reduction in the number of head of household with less than 4 years of schooling and a substantial increase in female headship between 1991 and 2000.

4.17 The microdata also allow us to look at the heterogeneity of impacts. In other words, one would like to test the hypothesis that the program had a symmetrical impact over all income and demographic groups. This heterogeneity can reflect both the ability of different groups to internalize benefits, as well as possible capture of benefits by select groups. One cannot a priori assess what would lead one group over another group to exhibit more or less benefits from the program. *Ex ante*, the literature on human capital and education points to the impact of schooling on the health production function. Therefore, it is possible that the more educated would be in a better position to make use of the beneneficial impacts of public goods provision. On the other hand, it is also possible that the more educated are better able to mitigate against not having goods like, for example, basic sanitation. The same argument can be made with respect to income. Additionally, given that there is substantial heterogeneity in income even within *favelas* (see introduction), looking at the differential impact by income quantile is of interest in order to answer wider questions of targeting and equity.

[56] Is is worth noting, however, that we are not controlling for other counfunding factors.

Table 4.8: Heterogeneity of impact, income quartiles, nearest neighbor method

Full Sample	Water	Sewer	Rubbish	Illumination	Illiteracy	Population
Quartile 1						
Diff-in-diff	0.081 [0.044]**	0.177 [0.053]+	0.060 [0.037]**	0.022 [0.029]	-0.158 [0.040]+	0.365 [0.122]+
Att 2000	0.072 [0.040]**	0.170 [0.039]+	0.040 [0.023]**	0.000 [0.002]	-0.110 [0.030]+	0.465 [0.092]+
Att 1991	-0.009 [0.021]	-0.007 [0.037]	-0.020 [0.031]	-0.022 [0.029]	-0.110 [0.030]+	0.100 [0.080]
Quartile 2						
Diff-in-diff	0.050 [0.037]	0.106 [0.050]*	0.035 [0.028]	-0.015 [0.028]	0.018 [0.039]	0.660 [0.134]+
Att 2000	0.055 [0.033]**	0.101 [0.038]+	0.014 [0.018]	0.000 [0.000]	0.028 [0.025]	0.637 [0.099]+
Att 1991	0.006 [0.017]	-0.006 [0.034]	-0.021 [0.025]	0.015 [0.028]	0.028 [0.025]	-0.023 [0.077]
Quartile 3						
Diff-in-diff	0.070 [0.041]**	0.107 [0.054]*	0.014 [0.033]	0.031 [0.024]	0.041 [0.042]	0.818 [0.134]+
Att 2000	0.043 [0.035]	0.109 [0.037]+	0.020 [0.021]	0.000 [0.000]	0.020 [0.027]	0.754 [0.095]+
Att 1991	-0.027 [0.018]	0.002 [0.035]	0.006 [0.025]	-0.031 [0.024]	0.020 [0.027]	-0.064 [0.092]
Quartile 4						
Diff-in-diff	0.017 [0.036]	0.079 [0.049]	0.036 [0.028]	0.011 [0.028]	0.029 [0.032]	0.598 [0.147]+
Att 2000	-0.002 [0.032]	0.037 [0.034]	0.010 [0.012]	0.000 [0.000]	0.025 [0.024]	0.700 [0.118]+
Att 1991	-0.019 [0.020]	-0.042 [0.035]	-0.027 [0.024]	-0.011 [0.028]	0.025 [0.024]	0.102 [0.090]
Eligible size	Water	Sewer	Rubbish	Illumination	Illiteracy	Population
Quartile 1						
Diff-in-diff	0.046 [0.050]	0.297 [0.065]+	0.042 [0.042]	-0.030 [0.040]	-0.104 [0.056]**	0.242 [0.113]*
Att 2000	0.076 [0.045]**	0.267 [0.047]+	0.023 [0.024]	0.002 [0.002]	-0.048 [0.042]	0.243 [0.095]+
Att 1991	0.030 [0.017]**	-0.030 [0.047]	-0.019 [0.034]	0.031 [0.040]	-0.048 [0.042]	0.001 [0.068]
Quartile 2						
Diff-in-diff	-0.014 [0.043]	0.159 [0.060]+	-0.004 [0.039]	-0.046 [0.026]**	0.034 [0.042]	0.295 [0.129]*
Att 2000	0.032 [0.040]	0.146 [0.046]+	0.024 [0.016]	0.000 [0.002]	0.013 [0.034]	0.413 [0.102]+
Att 1991	0.046 [0.020]*	-0.014 [0.043]	0.028 [0.035]	0.046 [0.026]**	0.013 [0.034]	0.118 [0.073]
Quartile 3						
Diff-in-diff	0.048 [0.039]	0.155 [0.056]+	0.014 [0.031]	-0.088 [0.024]+	0.055 [0.035]	0.456 [0.118]+
Att 2000	0.078 [0.034]*	0.153 [0.045]+	0.012 [0.013]	0.000 [0.000]	0.018 [0.028]	0.499 [0.098]+
Att 1991	0.030 [0.019]	-0.002 [0.035]	-0.002 [0.028]	0.088 [0.024]+	0.018 [0.028]	0.043 [0.070]
Quartile 4						
Diff-in-diff	0.067 [0.041]	0.106 [0.049]*	0.016 [0.025]	-0.018 [0.027]	0.042 [0.037]	0.731 [0.122]+
Att 2000	0.082 [0.032]	0.096 [0.036]+	0.000 [0.012]	0.000 [0.000]	0.022 [0.027]	0.649 [0.106]+
Att 1991	0.014 [0.021]	-0.010 [0.037]	-0.016 [0.020]	0.018 0.027	0.022 [0.027]	-0.082 [0.060]

+ 1 percent significance; ** 5 percent significance; * 10 percent significance

4.18 Table 4.8 shows several patterns that were absent in the aggregate findings. Only nearest neighbor estimates are reported. There are significant differences among different estimation methods, however, they are not systemic. Only results that are robust are commented. With respect to the provision of public goods, we see two interesting findings. First, the magnitude of impacts in sewerage appears to be greater for the poorest quartile. This result is found both in the full sample and in the sample of eligible size communities. The difference in impact from quartile one to quartile four is on the order of two to one (full sample) and three to one (eligible size). The inter-quartile results for rubbish collection are similar. Here we see that among the higher quartiles the results is not significant, while for the poorest quartile it is strongly significant. This result was not present in the aggregate data. The results for water also follow this pattern, although in this case it is not robust across specifications. The second finding regards the expansion of population in program *favelas*. Although the income results in aggregate did not suggest any type of real economy impact, one can see that comparisons between *favelas* and other communities suggest that the segment that has most grown vis-à-vis

other communities have been the higher income quartile groups. This can be (possibly) because program *favelas* have attracted higher income residents due to the intervention at a higher rate than non-program communities. Another interesting result refers to the fall in illiteracy rates for the lower income quartile group. This result holds for both full sample and eligible size sample. A worrying result refers to access to illumination services which did not increase for any of the analyzed quartiles for the full sample and in the case of the eligible size sample displays show a negative performance *vis-à-vis* non-treated *favelas*.

C. Impact on Rents and Investment

4.19 One of the possible effects of an urbanization program is the increase in the value of properties. Access to water and proper sewer, as well as road access are features that affect the value of the property. Information from other surveys shows that 30% of the households carry out some refurbishment after PROAP I, therefore, one should see an increase in the value of the properties. As we do not have data on the value of the properties before and after the program, we have to rely on rental rate values before and after the program and from different sources. The 1991 census collected information on rents paid, but the 2000 census did not. Given this limitation we have to rely on information from one of two sources. The first is the PCBR survey, in which case impact is identified from the order in which communities entered the program (i.e. some communities had already finished their works when the survey took place whereas others had not started at the time of the survey or were still in the very beginning). The second source are stacked cross-sections of the PNAD, in which case we have information before and after impact (and during), but our comparison group then becomes other non-*favela* low income housing communities matched with the *favelas* in the PNADs. It is important to note that the comparison low-income communities are identified, as with the case of a propensity score, with a probit on the probability of the community being a *favela*: only communities in the common support are selected as treatments. Furthermore, given the high rates of inflation in 1992, which likely affected both relative prices and recall error of respondents, the initial year is taken as 1993. This also has the advantage of being closer to the initial implementation of the *Favela-Bairro* project, although it limits the degree of comparability with other data sources, such as the census estimates. The estimated regression, with only treatment and controls, can then be represented as follows:

$$y_i = a + b_Z Z + g I + b_I XI + d_T T + d_I TI + e \qquad (3.3)$$

where Z represents control covariates, I is a dummy variable for *favela* status, and T is a dummy representing post-program years of 1998 and 1999; the years prior to program implementation, 1993 and 1995, are omitted. The impact of the program on the treated between year 1993-1995 and 1998-1999 can then be represented as

$$E(Y_{it}^1 - Y_{it-k}^1 \mid P'(X), Z, I = 1) - E(Y_{it}^0 - Y_{it-k}^0 \mid P'(X), I = 0) = d_I \qquad (3.4)$$

where $P'(X)$ can be interpreted as a propensity score in which the bandwidth encompasses all of the control observations. It is a Nearest Neighbour estimator with an

infinite neighbourhood, conditional on the common support[57]—in other words, all of the control observations in the common support serve as a match for each of the *favela* communities.

Table 4.9: OLS regression of rental rates, using communities matched to *favelas*

	Observed rental rate (2)					
	Favela		Formal city		Difference	
	Estimate	p-value	Estimate	p-value	Estimate	p-value
Crowding	13.450	(0.05)	9.839	(0.17)	3.610	(0.53)
Number of rooms	14.852	(0.01)	31.660	(0.00)	-16.809	(0.24)
Piped city water service	41.119	(0.03)	-17.393	(0.24)	58.512	(0.02)
City sewerage	-2.100	(0.90)	3.956	(0.70)	-6.056	(0.61)
Garbage collection	9.498	(0.45)	46.215	(0.00)	-36.717	(0.39)
Concrete structure	16.576	(0.21)	21.742	(0.01)	-5.166	(0.87)
Neighborhood distance to central district(1)	-116.256	(0.00)	-26.474	(0.42)	-89.781	(0.02)
Neighborhood income level	0.026	(0.67)	0.272	(0.01)	-0.246	(0.16)
1998-1999 year dummy (after program)	63.500	(0.00)	110.910	(0.00)	-47.410	(0.00)
Observations	526					
R squared	0.437					
	Predicted rental rate					
1998-1999 year dummy (after program)	8.659	(0.06)	25.720	(0.00)	-17.061	(0.00)
Observations	4417					
R squared	0.013					

(1) Measured by proportion of commuters who live one hour or further from work.

(2) Rent is predicted based on coeficients from 1993-1995 cross-sections

4.20 It should be noted that the direction of effect is by no means evident. According to the standard textbook explanation on rental rates and property values, in a frictionless environment with no transaction costs or information asymmetries, the relationship between the quality adjusted rental rate and the quality adjusted property value is given by the following:

$$R_t(q) = P_t(q)\left[(1+i)(t_P + t_Y) + d_t + r_t - E_t(P_{t+1})\right] \quad (4.1)$$

where R is the rental rate, P is the property value, i is the interest rate, tp and ty are the tax rate on property and the marginal income tax rate, d represents the rate of depreciation, r represents the risk premium, and $E(P)$ the expected change in property values. If the program changes the expectations regarding future property values, then the corresponding rental rate may fall. That is, one expects the program to have to

[57] A "hard" common support of similar income was also included. In other words, in a first step communities with incomes outside those of favelas are excluded. In a second stage, those within 95% of the common support are retained for the analysis.

countervailing effects, and increase in rental rates due to increases in quality and amenities, $\frac{\partial R_t}{\partial q}$, is expected to be positive, while $\frac{\partial R_t}{\partial E_{t+1}(P)}$, the change due to expectations, may well be negative. To deal with this problem, we also estimate hedonic rent functions and predict rental rates based on pre-program parameters. In effect, this produces an housing quality index function.

4.21 The top panel of Table 4.9 shows the results from the observed real rental rates, before and after the program, and in *favelas* and in other low income communities in the common support. The baseline is from the 1993 and 1995 years. The negative difference in the post-program dummy variable between *favelas* and other low-income communities indicate *favelas* have not seen an appreciation in rental rates relative to the control group of other low income communities in the common support. If anything, the data show depreciation. If we then turn our attention to the bottom panel, where predicted rental rates are presented[58], we see that as in the case of observed rent, there is an increase in the post-program period versus the baseline period. However, again, the overall impact parameter is shows a negative sign, indicating that even as measured by the housing determinants of property value, there is no appreciation in treatment versus control groups.

4.22 However, given that this comparison is between *favela* and comparable non-*favela* groups, two complications arise. First, one cannot differentiate between impact attributed to *Favela-Bairro* and to other neighborhood improvement projects such as *Bairrinho* and *Grandes Favelas*. The specification also does not separate improved versus unimproved favleas, and therefore can be thought of as a crude measure of "intent to teat", rather than as an impact of the treatment on the treated. The second, and equally serious problem is that formal communities may not in fact be adequate controls, particularly if these are characterized by more complete land markets—in which one would expect prices to clear more quickly—or even if they were subject to different public service interventions, or even investments by the private sector that may generate positive externalities. Lastly, as mentioned above, it is not clear that neighborhood improvements in public goods would appreciate rental rates to the same extent as they appreciate property value. All these caveats notwithstanding, at the end of the day the absence of any indication of appreciation—in this specification—is a relevant finding.

4.23 Table 4.10 shows that the difference-in-difference analysis for the sample of communities that can be identified in both 1991 census and PCBR show no effect on the treated communities. Treated communities had lower rents in 1991 compared to non-treated ones, but the rents according to the information collected by PCBR rose in the same pace for both of them. This increase was also documented using the PNADs.

[58] Predicted rents are based on an OLS regression of rents on number of rooms and number of rooms squared, number of bedrooms and number of bedrooms squared, city water hookup, city sewerage hookup, type of garbage collection (durect and indirect, no collection omitted), type of roof construction ("laje" and "telha"were the only included categories, other types of construction constitute the omitted dummy), the presence of concrete walls (alvenaria), and the average neighborhood comute to work time.

Table 4.10 - Difference-in-Difference Estimates for Community rents (using the PCBR survey)

	coef	std. Err
rooms	0.145	[0.009]***
wall	0.503	[0.080]***
roof	-0.152	[0.140]
concentration	-0.006	[0.009]
toilet	-0.049	[0.081]
water	-0.314	[0.055]***
sewer	-0.014	[0.032]
rubbish	0.237	[0.044]***
year(2000)	0.681	[0.080]***
treated	-0.214	[0.067]***
treated*year(2000)	**0.051**	**[0.070]**
constant	3.863	[0.180]***
N	4875	
Adj.R-squared	0.231	

4.24 As with the case using the PNADs, this analysis has be taken only as exploratory, given that we were unable to measure the time that each of the communities was in the program, and given the relatively small size in the PCBR analysis. To deal with changes in relative prices, we deflated the reported rental value according to the month and year of reference of the survey. An additional complication is that all communities interviewed were due to receive the program sooner or later, and most of the treated communities still had sanitation projects under way when the survey was conducted. Therefore, the lack of an observed effect may be due to the very short period in which communities would have been in the program. Unfortunately the documentation obtained from the program is not detailed enough to offer a monthly figure as to the progress in execution of the various public works. For this reason even the classification between treated and comparison communities cannot be guaranteed with precision[59]. However, the fact that the results from the PNADs and the PCBR point in the same direction lend more credence to our findings. At the very least we can say that the data are consistent in that there is no evidence that *Favela-Bairro* has had a significant impact on housing values in *favelas*, beyond what is observed in both other favelas and comparable non-favela communities.

D. Impact on Mortality

4.25 One of the outcomes emphasized among the goals of the program was the reduction of the incidence of diseases caused by lack of sanitation. In order to assess the impact of the program on this issue we will rely on mortality data provided by the Municipal Health Department based on death certificates. By Brazilian law every death has to be documented in a form that collects information such as residential address, race, gender, primary death cause, etc. Such rich administrative record has been improving over time,

[59] We experiment with alternative classifications based on the poor information about work's progress and the results do not change

but even so the reported addresses do not allow us to identify, with a reasonable degree of certainty, the community where the dead lived[60]. The neighborhood where he lived, on the other hand, is very well documented. This allows for a tracking of death counts by death causes in each one of the 153 neighborhoods.

4.26 A second method used in the literature to assess the impact of greater water and sanitation services on health is infant mortality, defined as the probability of having a child under the age of one die during the reference period. This data is also available from the administrative records, and is analyzed in the same way as the data on causes that are documented to have been associated with vector-borne disease. This is done in part to account for cases of misdiagnosis, where the primary case of fatality may have been exacerbated by secondary parasitic or viral vector-born infections.

4.27 The literature on crime and violence suggests that improvements in the access to public services, as well as an increased presence of the state more broadly may also be associated with decreases in violent crime. To investigate this possibility in *favela-bairro*, we look at causes of death by homicide. Although the theoretical underpinnings are perhaps weaker than those relative to health and sanitation, they with a methodology analogous to that used in the data on child mortality, described below.

4.28 The dependent variable in our analysis is the (P_{ij}) proportion of death in group i (i = death caused by lack of sanitation related disease; homicide; children younger than 1 year) in neighborhood j over all deaths registered in that neighborhood. In order to capture how changes occurred between 1995/1996 and 2000/2001[61] and how it differed according to the percentage of household in treated communities that were in those neighborhoods, we put together data for 1995/1996[62] and 2000/2001 and model that death proportion according to different specifications. All specification include an interaction between the proportion of household treated in the neighborhood and a dummy for year 2000/2001 as the proportions did not change much; most of the change was due to variation between 1995/1996 and 2000/2001. The regressions were weighted by the total number of deaths in each neighborhood[63]. The empirical specification can then be represented as follows:

$$P_{ij} = a + b X_j + g\,2000_j + h P_j^{FB} + d\,2000_j P_j^{FB} + e_j \qquad (4.1)$$

where the impact parameter is given by $E\left(\frac{\partial P_i}{\partial P^{FB}} \mid t = 2000\right) - E\left(\frac{\partial P_i}{\partial P^{FB}} \mid t = 1996\right) = d$.

4.29 Table 4.11 presents results of cause of death due to diseases related to lack of sanitation. The sign and significance of most of the variables are as expected: connection to the city

[60] This fact may me due to some "stigma" for living in communities leading to misreport or due to problems when recording the manuscript information into electronic format.

[61] We aggregate the 1995 and 1996 deaths and the 2000 and 2001 deaths so that we could increase the number of observations.

[62] Actually the data on water and sewer access, rubbish collection, income and schooling come from the 1991 census, because the 1996 Population Counting did not have detailed information on those variables. The information on death, however, are based on 1996 data.

[63] We also experiment weighting by the total population of the neighbourhood, but the results were basically unaltered.

sewerage service is associated with lower mortality rates, as is a more schooled household head. Surprisingly we did not find that the water access impacts the incidence of these mortality rates for most specifications, nor did the household income level in one of the specifications[64]. In terms of the program's impact, the proportion of treated households does not decrease the mortality rate over time. The proportion of diseases related to lack of sanitation slightly decreased between 1995/1996 and 2000/2001, but this fall was not significantly higher in neighborhoods with a higher proportion of treated communities, despite the negative sign.

Table 4.11 – D-in-D estimates for the Proportion of Deaths caused by diseases related to lack of sanitation

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
year==2000	-0.002	-0.002	-0.002	-0.003	-0.002	-0.002	-0.002
	[7.92]**	[8.01]**	[7.92]**	[8.25]**	[6.45]**	[5.90]**	[6.44]**
prop.treated	0.015	0.013	0.015	0.015	0.015	0.014	0.014
	[2.46]*	[2.16]*	[2.39]*	[2.44]*	[2.37]*	[2.27]*	[2.33]*
(year==2000)*prop.treated	-0.007	-0.007	-0.007	-0.007	-0.007	-0.007	-0.007
	[1.07]	[1.09]	[1.04]	[1.00]	[0.99]	[1.03]	[1.13]
proportion of HH in favelas		0.002					
		[3.06]**					
water dummy			-0.001	0.003	0.006	0.005	0.007
			[0.41]	[1.03]	[1.91]	[1.38]	[2.08]*
sewerage dummy				-0.002	-0.001	-0.001	-0.000
				[2.65]**	[1.56]	[0.61]	[0.17]
garbage dummy					-0.011	-0.008	-0.008
					[2.08]*	[1.62]	[1.58]
log HH income						-0.001	0.002
						[2.81]**	[2.21]*
average yrs schooling							-0.001
							[3.04]**
Constant	0.004	0.004	0.006	0.003	0.010	0.013	-0.002
	[18.55]**	[15.48]**	[2.04]*	[1.27]	[2.31]*	[3.14]**	[0.33]
Observations	306	306	306	306	306	306	306
R-squared	0.28	0.30	0.28	0.30	0.31	0.32	0.34

Robust t statistics in brackets
* significant at 5%; ** significant at 1%

4.30 Interestingly, the proportion of treated slums in the basic specification in column [1] is significant, even when we introduce the proportion of households in all communities (column[2]), it does not lose significance. This suggests that that targeting of urbanization program in *favelas* seems to be a correct in the sense that these communities have greater needs. However, as commented in the last paragraph the interaction between the year after-treatment and the proportion of treated households in the neighbourhood is never significant, despite being negative. This suggests that the health and sanitation components of *favela-bairro* have, surprisingly, no impact on mortality associated with vector-born disease. A possible explanation for this puzzle may be in the fact that most sanitation services were already relatively high in *favelas*, and that the income levels (and

[64] The positive and statistical significant estimate of the coefficient for income in specification [7] may be due to a high correlation between income and schooling causing some multicolinearity.

overall levels of development) of *favela* residents may have been high enough so that deficits in sanitation would not be a factor in mortality. Along the same lines, it is possible that mortality by vector-bourn disease is such a rare event, that most of the benefits of increased sanitation may be reflected in child morbidity and development rather than in mortality. Given that data on these variables was not collected before 2000 in the treated and control communities, one can only conjecture as to the development impact of the program along this dimension.

4.31 A second method used to pick up the effect of sanitation is to look at the incidence of mortality under the age of one. Table 4.12 shows that infant mortality fell between 1991 and 2000, but this fall was not larger in neighborhoods with a higher proportion of treated communities; on the contrary, in most specifications it seems that there is a positive non-significant effect. However, the endogeneity of program placement could be the culprit, i.e, the selection of communities with higher chances of fatality, may have been driving the results. To the extent that this bias is stationary it would be, of course, differenced out in the double difference estimator, which as mentioned above is not significantly different from zero. Most other results in the table are similar to those of cause of death due to vector-bourn disease. But it is worth to notice that strong impact that connection to the city sewerage system has on bringing down child mortality. Income and schooling also seem to have a negative impact on child mortality.[65] Nevertheless, previous researches found very weak links between child mortality and socio-economical conditions in Rio de Janeiro. Campos et al. (2000) did not find a significant correlation between child mortality and living conditions while analyzing borough's pattern of child mortality. Similarly, Szwarcwald et al. (1999) and Szwarcwald et al. (2002) do not find robust evidence that child mortality may be correlated with the proportion of population living in favelas, once one control for porverty and inequality index. The general conclusion of these papers is that the lack of significant spatial concentration of child mortalilty indicates that public policy should focuses on improving the health services for deprived communities rather than expecting that infrastructure upgrading would reduce child mortality on its own. In support to this recommendation, Sastry (1996) did not find significant effects on child mortality of improvements in community sanitation, either in the Northeast or in the South/Southeast of Brazil. Moreover the measures of community infrastructure – water supply, sanitation, electricity, and the presence of a rubbish collection and public cleaning service – were significantly and negatively correlated with child mortality risks only in the Northeast region. He argues that there is evidence that after a certain threshold substitutability between maternal education and community services may change to complementarities as these services become more prevalent. This would be precisely the case of the Southeast, in general, and of Rio de Janeiro in particular. In terms of public policy, he argues that larger reduction in mortality in the Southeast could be achieved to exposure to messages on child care and family planning.

[65] Again, possibly multicolinearity between schooling and income may it hard to understand its coefficient when put together in the model.

Table 4.12 – D-in-D estimates for the Proportion of Deaths of Child younger that 1 year

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
year==2000	-0.014	-0.014	-0.013	-0.021	-0.015	-0.011	-0.014
	[4.22]**	[4.39]**	[4.17]**	[6.96]**	[5.04]**	[4.42]**	[5.00]**
prop.treated	0.016	-0.014	0.007	0.014	0.014	-0.004	-0.009
	[1.03]	[0.85]	[0.46]	[0.94]	[0.99]	[0.30]	[0.50]
(year==2000)*prop.treate	0.021	0.021	0.029	0.037	0.036	0.030	0.025
	[0.91]	[0.86]	[1.23]	[1.72]	[1.75]	[1.45]	[1.07]
proportion of HH in favelas		0.036					
		[5.66]**					
water dummy			-0.077	0.030	0.069	0.026	0.063
			[2.41]*	[1.07]	[2.54]*	[1.01]	[2.11]*
sewerage dummy				-0.055	-0.044	-0.025	-0.020
				[5.49]**	[4.90]**	[3.32]**	[2.91]**
garbage dummy					-0.135	-0.071	-0.068
					[2.97]**	[1.91]	[1.97]*
log HH income						-0.016	0.016
						[7.09]**	[1.57]
average yrs schooling							-0.010
							[3.27]**
Constant	0.043	0.038	0.119	0.064	0.145	0.215	0.030
	[17.11]**	[12.99]**	[3.76]**	[2.72]**	[3.51]**	[5.95]**	[0.42]
Observations	306	306	306	306	306	306	306
R-squared	0.12	0.22	0.15	0.31	0.36	0.52	0.57

Robust t statistics in brackets
* significant at 5%; ** significant at 1%

4.32 Although a reduction in the incidence of homicide was never an intended program impact, the literature suggests that increasing the presence of the state, as well as increases in public services have been associated with reductions in violent crime. Table 4.13 shows the impact results for homicides. As in the cases above, the estimates do not show a reduction in homicides greater in neighborhoods with more interventions communities. Although the sign of the impact is negative throughout specifications, it is never significant at conventional levels. We see that all of the socioeconomic variables—schooling, income—are significant. Higher schooled, older and wealthier neighborhoods can expect lower homicide rates[66]. We also see that the presence of the provision of state services also is associated with decreases in homicide rates[67].

[66] Again it seems that there is some multicolinearity going on between age profile and proportion of household in slums.
[67] Szwarcwald et al. (1999) also found that the homicide rate is highly correlated with that proportion of household in favelas.

Table 4.13 - Difference in Difference estimates for the Proportion of Homicides

	[1]	[2]	[3]	[4]	[5]	[6]
year==200	-0.005	-0.005	-0.005	-0.017	-0.002	-0.004
	[1.03]	[1.08]	[1.00]	[3.83]**	[0.81]	[1.32]
prop.treated	0.037	-0.010	0.035	0.046	0.007	0.005
	[3.40]**	[0.70]	[3.36]**	[4.25]**	[0.84]	[0.54]
(year==2000)*prop.treate	-0.004	-0.005	-0.002	0.011	-0.003	-0.005
	[0.14]	[0.16]	[0.06]	[0.46]	[0.14]	[0.25]
proportion of HH in favelas		0.056				
		[4.88]**				
water			-0.021	0.147	0.096	0.117
			[0.62]	[4.44]**	[3.72]**	[4.17]**
sewerage				-0.087	-0.036	-0.033
				[6.31]**	[4.52]**	[4.26]**
log HH					-0.032	-0.015
					[14.82]**	[1.64]
average yrs						-0.005
						[2.11]*
Constant	0.058	0.050	0.078	-0.007	0.219	0.117
	[17.41]**	[13.16]**	[2.35]*	[0.25]	[7.96]**	[2.03]*
Observation	306	306	306	306	306	306
R-	0.02	0.13	0.02	0.22	0.59	0.60

Robust t statistics in
* significant at 5%; ** significant at

4.33 Overall, these results suggest the program had little effect on the causes of death related to vector-borne disease and their primary victims—children. However, as mentioned previously, the empirical specification used may be confounding the effect of program selection with the actual program impact, at least to the extent that characteristics related to both selection and cause of death is not stationary, and therefore cannot be netted out in a double-difference. Although the structural homicide model panned out many of our *ex ante* expectations, here too we see an absence of any measurable impact due to the program.

V. CONCLUSIONS

5.1 The analysis above is an assessment of how the *Favela-Bairro* program has changed the quality of life of the residents of the low-income neighborhoods intervened by the project. There are three fundamental conclusions that fall out of the analysis. The first conclusion is that given the results obtained by both the propensity score matching and by the exploiting the longitudinal nature of the household surveys available, the evidence of a credible impact of *Favela-Bairro* has been on the program's outputs. During the decade we see a substantial increase in the coverage of water and rubbish collection in *favelas* that outpaced the comparison groups identified. The impacts on sewerage was the most significant one in the aggregate level, moreover, an analysis by income quartile reveals that the poorest quartiles did benefit from sewerage, while the richest quartiles did

not, or did in a lesser extent. This result—by quartile—is also seen with respect to water, rubbish collection and illiteracy.

5.2 In both the value of property, as measured by rent, and the in most of the health outcomes, however, the data have not produced evidence that the program has significantly improved the quality of life of residents. With respect to rent, the results of observed rental rates, the reflexive evaluation shows an increase in rents, but this was actually a citywide event; the impact parameter shows a decline in *favelas* relative to other low-income communities. Whereas the impact on observed rental rates was negative. This is true with respect to predicted rents also, which is a measure of housing quality. These results come with the caveats already mentioned in the text, that the control groups for rents may not be properly matched to the treatment group in 1995, given that the control was low-income but not *favelas*. In terms of earnings, the results also do not suggest that the program was able to generate any real economy impacts. It should be noted that this impact was in fact never anticipated by the project's design.

5.3 The second result is that we see a marked difference in the characterization of the project's impact when no control group is used. In terms of results a reflexive comparison that does not take into account a counterfactual leads to a super-estimation of project benefits that varies by output and outcome analyzed but that in all cases is very significant. In the case of sewer coverage and rubbish collection this over-estimation is in the neighborhood of 100%. In the case of housing values, an assessment of the *favelas* based on a simple before-and-after comparison shows an impact of between 30 and 60 percent, depending on the specification; yet this impact is dwarfed by the overall appreciation of housing values in the city over the second half of the 1990s. The natural corollary is that in non-stationary scenarios reflexive before and after estimations of impact may grossly misrepresent the project or program impact, possibly leading to erroneous and costly policy recommendations.

5.4 The last result is perhaps the most important, in part because it qualifies most of the findings of this evaluation. Given the high degree of geographic targeting, yet the relatively small size of targeted units, the data available was inadequate to answer many of the fundamental evaluative questions raised in the introduction and throughout the paper. The household survey data, even when stacked over multiple cross-sections, produced sample sizes that were small. Furthermore, since only a small sample of census tracks are actually sampled, the sample of treated—even when one considers all *favelas*—was of only 30 census tracks. On the other hand, approaches that used census data were likewise problematic, since (i) the census does not contain useful information on many of the outcomes of interest, and (ii) the relatively large time elapsed between census years necessarily implies that project impact will be confounded with other macro and regional effects that occurred between the first census and actual implementation. In short, the nature of the program called for a specialized sample survey structure in which communities both in and outside the project could be canvasses—preferably based on an assignment criteria uncorrelated with the outcomes of interest. Parts of this approach were pursued at different points in the project's execution, producing a varied collection of different datasets on beneficiaries. However, these were mostly not comparable, and

in any case did not include information on non-participants. These characteristics of the data limit the quality of comparisons over time of the same communities and make comparisons over time between program and non-program communities impossible. This implies that beyond expansion of public services, one is severely limited regarding what can the impact of *Favela-Bairro* on key development outcomes.

REFERENCES

Alonso W. *Location and Land Use*, 1964, Harvard University Press: Cambridge, MA.

Baum-Snow, N. "The Effects of Changes in the Transportation Infrastructure on Suburbanization: Evidence from the Construction of the Interstate Highway System", mimeo, 2004.

Besley, Tim "Property Rights and Investment Incentives: Theory and Evidence from Ghana". *Journal of Political Economy* 1995, 103(5): 903-937.

Bénabou R. "Workings of a city: Location, education and production", *Quarterly Journal of Economics*, 1993, 108 (3): 619-652

Borjas, George J., "Ethnicity, Neighborhoods, and Human-Capital Externalities" *American Economic Review*, June 1995, 85 (3): 365-390.

Brueckner J. K., J.-F. Thisse, and Y. Zenou "Why Is Central Paris Rich and Downtown Detroit Poor? An Amenity Based Theory", *European Economic Review*, 1999, 43 (1), 91-107

Campos, T. P., Carvalho, M. S. and Barcellos, C. C. "Mortalidade infantil no Rio de Janeiro, Brasil: áreas de risco e trajetória dos pacientes até os serviços de saúde". *Revista Panamericana de Salud Publica*. 2000, 8 (3): 164-171.

Case, Anne and Lawrence Katz, "The Company you Keep: The Effect of Family and Neighborhood on Disadvantaged Youth", NBER Working Paper 3705, May 1991.

Cutler, D.M. and E.L. Glaeser, "Are Ghettos Good or Bad?" *Quarterly Journal of Economics*, 1997, 112: 827-72.

Cutler, D., E. Glaeser, and J. Vigdor, "The Rise and Decline of The American Ghetto." *Journal of Political Economy*, 1999, 107: 455-506.

Cutler, D., E. Glaeser, and J. Vigdor, "Ghetos and the Transmission of Ethnic Capital." *mimeo*, 2002.

de Soto, Hernando. *The Other Path.* 1989, Harper and Row, Publishers.

Ellickson, R.C. *Order without Law: How Neighbors Settle Disputes*. 1991, Harvard University Press, Cambridge, MA

Field, Erica. "Property Rights and Household Time Allocation in Urban Squatter Communities", mimeo, 2003.

George Galster, "Assessing the Causes of Residential Segregation: A Methodological Critique," *Journal of Urban Affairs*, 1988, 10: 395-407.

Glaeser, Edward, Sacerdote, Bruce and Jose Scheinkman, "Crime and Social Interactions" *Quarterly Journal of Economics*, May 1996, 111 (2): 507-548.

Glaeser E. L., M. E. Kahn, and J. Rappoport "Why do the poor live in cities?" NBER Working Paper 7636, 2000.

Jimenez, Emmanuel. "Tenure Security and Urban Squatting" *Review of Economics and Statistics*, 1984, 66 (4): 556-567.

Kanazawa, Mark. "Possession is Nine Points of the Law: The Political Economy of Early Public Land Disposal," *Explorations in Economic History*. 1996, 33: 227-49.

Lanjouw, J. O. and Philip Levy "Untitled: A Study of Formal and Informal Property Rights in Urban Ecuador" *The Economic Journal*, 2002 112(482): 986-1019.

Manski, C. F. "Learning About Treatment Effects from Experiments with Random Assignment of Treatments." *Journal of Human Resources*, 1996, 31(4) p. 707-733.

Mills E. S. "An aggregative model of resource allocation in a metropolitan area", *American Economic Review Papers and Proceedings*, 1967, 57 (2), 197-210.

Möbius, Markus. "The Formation of Ghettos as a Local Interaction Phenomenon" 2000, mimeo.

Muth R. F. *Cities and Housing*, 1969, University of Chicago Press: Chicago

Morais, M.P., Cruz, B.O, and C. Oliveira, "Residential Segregation and Social Exclusion in Brazilian Housing Markets", IPEA TD, 2003.

Olinto, Pedro and Michael Carter. "Getting Institutions 'Right' for Whom: Credit Constraints and the Impact of Property Rights on the Quantity and Composition of Investment" mimeo, 2003.

Rosenbaum, PR, and DB Rubin, "The central role of the propensity score in Observational studies for Causal Effect", *Biometrika*, 1983, Vol. 70, 41-55.

Sastry, Narayan (1996) "Community Characteristics, Individual and Household Attibutes, and Child Survival in Brazil". *Demography*, 1996, 33 (2): 211-229.

Schelling, T. C "A Process of Residential Segregation: Neighborhood Tipping," in *Racial Discrimination in Economic Life*, ed. by A. Pascal. 1972, Lexington Books, Lexington, MA.

Szwarcwald, C. L., Bastos, F. I. , Esteves, M. A. P., Andrade, C. L. T, Paez, M. S., Medici, E. v., and Derrico, M. "Desigualdade de Renda e Situacao de Saude: o Caso do Rio de Janeiro" *Cadernos de Saude Publica*, Rio de Janeiro, 1999, 15 (1): 15-28

Szwarcwald, C. L., Andrade, C. L. T and Bastos, F. I. "Income Inequality, residential poverty clustering and infant mortality: a study in Rio de Janeiro, Brazil" *Social Science & Medicine,* 2002, 55: 2083-2092.

Vianna, Andrea. "As Políticas de Habitação no Rio de Janeiro e Seus Resultados", mimeo, 2004.

Wheaton, W. "Income and Urban Residence: An Analysis of Consumer Demand for Location" *American Economic Review*, 1977, 67 (4): 620-631.

Wilson, W. J. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. 1987, University of Chicago Press, Chicago, IL.

World Bank "Brazil. Progressive Low-Income Housing: Alternatives for the Poor", December 2002.