Location of Economic Agents in Brazil: An Empirical Investigation

Filipe Lage de Sousa[†]

The London School of Economics and Political Science

Thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (Ph.D)

July 2009, London

† My most heartfelt thanks go out to my entire family who has always supported me during my PhD. The author is especially grateful to Henry G. Overman for his helpful and wise comments and suggestions. Many people have read parts of my thesis and contributed with excellent comments, for which I am extremely grateful as well. I also thank all those who have provided me access to any data used in this work. This research is supported by BNDES and by the Programme AlBan, the European Union Programme of High Level Scholarships for Latin America, scholarship number E04D029558BR, to which I am also grateful. The usual disclaimer applies.

Abstract

This study focuses on regional aspects of the Brazilian economy. Three avenues are explored: two related to individuals' and firms' location decision and the last regarding public policy. In general, firms and individuals seem to be moving away from the main economic centre in Brazil, Sao Paulo. It seems nevertheless that public policy is not very related to these movements. In other words, government interventions to accelerate growth in less developed regions have not achieved their goal.

Chapter 2 addresses the issue of internal migration in Brazil. This chapter investigates the influence of amenities and/or disamenities on migration flows, which is an issue not yet fully covered by the literature. It investigates whether changing dwellings across cities is associated specifically with violence using urban-urban migration data at municipality level. Results show that migration is affected by violence not only locally, but also when neighbouring effects are taken into account. These findings back up previous research which evidenced an inverse relationship between city size and violence.

Turning to firms, Chapter 3 explores the role of geography in the location of manufacturing and of regional disparities in wages. According to theoretical models, employment concentrates closer to the market when increasing returns to scale are taken into account. As a consequence, regional wages are a decreasing function of transport costs to markets, since firms tend to compensate for these costs by paying less to their employees. Trade shocks may impact these regional wage disparities by making foreign markets relatively more attractive for firms than internal markets, or vice-versa. This chapter tested these hypotheses using Brazilian regional data. Having two isolated trade shocks, Brazil provides an excellent case for testing which shock was more effective in reducing regional disparities. Results show that regions with higher transport costs tend to have lower wages and a reduction in this cost through trade shocks has affected these

regional disparities. However, it is not possible to distinguish which trade shock was more efficient to impact these regional unbalances.

Chapter 4 evaluates the effects of the Brazilian Development Bank (BNDES) loans on firms' productivity. The importance of BNDES in the Brazilian economy is quite sizeable, reaching over 10% of aggregate investment. Using micro level data, it was possible to investigate the impact on productivity, but also distinguish its effects between large and small projects as well as between rich and poor regions, since regional development is one of its statutory goals. Results suggest BNDES loans have no effect on firms' productivity, even though some association was found without controlling for all firms' characteristics.

Overall, some lessons may be learned after this work. Not only are economic reasons key determinants for individuals' and firms' location decision as shown in Chapters 2 and 3 but also some other factors seem to be important as well. Social amenities, locally and in surrounding areas, are highly correlated to individuals' migration decisions in the Brazilian case, especially violence. For firms, economic reasons prevail since trade shocks appear to change regions' attractiveness between internal and external market. Last, but not least, government intervention does not seem to be associated to firms' productivity after BNDES loans.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Regional disparities exist worldwide. Although geographic inequalities may persist, the relevance of each region might change over time. Individuals' and firms' location decision may be considered one important reason to explain those changes across regions. This thesis contributes to the literature of regional economics by exploring agents' location decision, which might culminate in strengthening or weakening regional differences.

In order to investigate those issues, three different agents are analyzed here. First, individuals are essential to understanding the distribution of settlements in space. What influences individuals' decision on whether to migrate or not should be explored. On the other hand, firms' strategy on where they maximize their profits plays an important role since people may migrate towards places with job opportunities. Government interventions may also be relevant as some policies try to promote regional development in less developed areas.

One of the issues investigated by the literature of regional economics is where individuals decide to locate. Indeed, their objective is to maximize their utilities. Having more benefits than costs is not a sufficient condition for deciding where to live. People will live only where the difference between benefits and costs achieves its peak. Whenever individuals perceive that one locality might provide them the highest utility, they feel tempted to move and some eventually migrate. Those migration decisions affect regional unbalances; therefore they are relevant to understanding why regions differ. Chapter 2 investigates which factors are associated with migration flows.

Migration is not a new issue. Actually, it has occurred since the beginning of human life on earth. Just as they did in the very remote past, individuals move from one place to another in order to maximize their utility. Many factors interfere in their migration decision and they vary from economic, social and environmental aspects. Not only factors are covered by the migration literature but also migrants' profile, when they move, where they go, and so forth. There are different types of migration as well; some are for example permanent while others only transitory. Internal migration exists as well as international. In summary, there is a diverse range of classifications and issues which the migration literature covers.

As a worldwide phenomenon, migration occurs in rich and poor countries alike, or even between them. Regarding only internal migration, countries from any classification have similar trends and patterns, but differences also exist. For example, dealing only with formal jobs in the developing world is not sufficient, since the informal sector plays an important role. In the developed world, access to housing plays a more substantial barrier than in poor countries, since low cost and precarious housing projects are easily constructed in shanty towns. The reasons for changing dwellings may slightly differ between developed and developing world, but they are essentially very similar.

Focusing on internal migration, Greenwood (1997) provides a survey of migration within borders of developed countries while Lucas (1997) of the developing world. One crucial distinction of these two surveys is which type of origin and destination, whether urban or rural, is mostly investigated in each case. While Greenwood lists much more urban-urban migration works in the developed world, Lucas shows more evidence of rural-urban migration in the developing world. Brazil hence shows a much more different pattern in which urban-urban migration flows are much more relevant since their importance has been increasing over the years.

One common issue relating these two surveys is the focus on economic factors for explaining migration decisions. Most of the migration literature emphasizes the economic reasons for why people move from one place to another. Although economic aspects are substantially relevant for migration decisions, some other issues might also be important, such as amenities.

There are two types of amenities: natural and social. While the former is essentially defined by physical attributes, the latter is a result of social interactions among inhabitants in every locality. One part of the literature, initiated by Rosen (1979), has evaluated whether these amenities may explain regional differences in economic variables, such as wages and housing prices. If they do, then differences in wages and housing prices are essentially the main part in explaining migration decisions. Results bear out however that basically natural amenities present robust results to explain these regional inequalities. Social amenities nevertheless do not show robust outcomes. These outcomes are corroborated by other papers in the literature.[1] Due to these evidences, Chapter 2 investigates the relationship between Brazilian urban-urban migration flows and social amenities after controlling for many other cities' attributes.

Among all social amenities, violence deserves closer attention for various reasons. First, it is a crucial problem in urban areas as Habitat (2007) states, especially in the developing world. Therefore, its importance has been increasing over the years. Second, crime rates are spatially concentrated in metropolitan areas in Brazil which have reduced their participation in urban population recently. Third, spatial effects in neighbouring areas may occur due to social interactions between individuals in contiguous areas. In order to show the importance of violence in urban aspects of the Brazilian society, the federal government has launched a national plan to reduce criminality in the 11 most violent

[1] A detailed discussion on this issue is presented in Chapter 2.

metropolitan areas, named PRONASCI.[2] Considering these arguments, Chapter 2 uses migration flows between Brazilian cities from 1985 to 2000 to evaluate whether they are correlated to social amenities, especially violence.

In order to evaluate this relationship, the econometric strategy explores basically two different gravity equations: a restricted form focusing on destination attributes; an unrestricted form using both origin and destination characteristics. This empiric methodology is extensively used by the migration literature. The main reason lies in the fact that regions have pushing and pulling factors which attract or repel inhabitants. The philosophy behind this methodological approach is which attributes are actually correlated to migration flows.

Aside from the traditional specification, different approaches are also tested, such as short- versus long-distance migration flows. Investigation of violence spatial lag effect is carried out as well in order to examine whether violence in migration flows may be correlated to it in neighbouring areas, since criminality affects not only where it occurs but also surrounding areas, as mentioned earlier.

Overall, results bear out that some economic variables are correlated to migration flows, such as services wages (nominal and real), housing prices, *inter alia*. With regards to amenities, outcomes suggest that they are intrinsically correlated to migration flows. A proxy for entertainment options has shown that people migrate towards places with higher values of this amenity. Population density, used as a proxy for a social disamenity, also presents the expected sign, which is that people are avoiding destinations which are densely populated. As it regards violence, not only the crime measure used shows a negative association with migration flows, but also other types of crime captured by population size.

[2] More details can be found at http://www.mj.gov.br/pronasci/.

By investigating spatial effects, results suggest their negative correlation as well. It is therefore safe to conclude that social amenities are thus relevant for migration decisions.

Individuals decide to live where opportunities to work arise, which is strongly related to firms' decision on where they locate. Just as with individuals, firms also decide where to locate their production plant in order to maximize their objective goal. The only difference here is that the objective goal changes from utility towards profits where the benefits are represented by revenues and costs are self explained. One branch of the literature deals with firms' location decision (New Economic Geography - NEG), where firms tend to locate close to the market in order to face lower transport cost to ship their goods to consumers. As a consequence, wages are higher close to markets. A negative correlation exists therefore between wages and transport cost. The economic intuition behind this argument is that firms facing higher transport cost need to offer lower wages in order to be competitive in the market. This trade-off between wages and transport cost faced by firms results in regional wage inequalities.

According to NEG literature, a reduction in trade costs may affect the regional distribution of economic activity within a country, and therefore regional wage disparities might change. Using Brazilian regional data, Chapter 3 investigates whether wages may be explained by demand linkages and indeed if trade shocks are able to interfere in this relation. Brazil seems to be a good example to investigate these issues because it has experienced some unexpected trade shocks which have affected their economy.

One shock occurred in the beginning of the 1990s: unilateral liberalization process. During this period, the weighted average nominal import tariff reduced from nearly 40% in 1988 towards just above 10% in 1994. This drastic reduction without any previous announcement from the new democratic president elected in 1989 increased substantially the competition in the Brazilian economy. Imports more than doubled from 1990 to 1995. It

started at around US$ 20 billion in the beginning to reach nearly US$ 50 billion in 1995. This trade shock may have a substantial effect on firms' location decision, since an increase in competition with imports may give incentives for firms to locate far from foreign market.

The expansion shock happened in 1999, when there was a real exchange rate devaluation of nearly 50% in a single year. Any Brazilian product became 50% more competitive after just one year. All trade goods were benefited by this unexpected trade shock. Therefore, exports boomed. They grew from just below US$ 50 billion in 1999 to nearly US$ 100 billion in 2004. In geographical terms, regions close to the external market became more attractive for firms' location after the exchange rate devaluation. These two shocks have some particularities which differentiate from a trade agreement, the most famous trade shock mentioned in the trade literature. First, both shocks experienced by the Brazilian economy were more unexpected than a trade agreement. Trade agreements generally take time to reach a final deal when trade costs are eventually reduced. Therefore, firms knowing the development of the generally long negotiations might start to change their strategy before the agreement is reached. A trade agreement hence may not be necessarily a shock, since firms may anticipate their actions, including location terms. Second, pushing (from imports) and pulling (from exports) factors in regions close to the external market occur simultaneously in a trade agreement. In the Brazilian case, they occurred at different periods. It is therefore possible to evaluate which force may influence regional disparities more substantially. It is safe to conclude that these Brazilian trade shocks seem to be not only more adequate than a trade agreement but also richer in investigating regional effects of a trade policy. Using data from 1985 to 2004, a period which encompasses these two trade shocks, Chapter 3 investigates these issues as well.

A detailed disaggregated regional panel data is used for testing these hypotheses mentioned earlier. As regional wages disparities may be explained by demand linkages,

two ways are explored by the literature: either by transport cost; or by market potential. First, a panel data approach with over 500 regional units is used to test whether wages may be explained by the transport cost to internal and external market. The background explanation of using panel approach lies in the fact that any unobservable regional attribute is included by using this method. In order to test whether trade shocks are able to affect this relationship between transport cost and wages, the empirical specification evaluates whether this association has remained stable over time. If not, then which market (internal or external) was affected and after which shock (reducing export or import cost).

One shortcoming of using transport cost is the idea that only few places may represent either internal or external markets. One way of tackling this issue is by using a different measure: market potential itself. The recent literature has deeply explored this venue, even though this measure reveals other drawbacks: endogeneity. However, a variety of instruments are used to overcome this problem by the literature. In this paper, two instruments are used: population size; and distance to economic centres. Therefore, another panel data approach is estimated by using market potential as explanatory variables for regional wages. The philosophy of using the panel data approach as well as how they are implemented are exactly the same as those used with transport costs.

Testing the hypothesis of whether demand linkages are relevant to explain regional wage disparities is important for the literature of regional economics. Knowing how important markets are for explaining differences in remuneration helps to understand the agglomeration and dispersion forces in economic geography. Additionally, evaluating whether trade shocks might influence these inequalities seems to be important for the international trade literature, since results have not detected any interference so far, as shown in Hanson (1996; 1997). Aside from academic interests, these issues are also relevant for policy makers, since knowing how the economic geographical shape might

change after a trade policy helps government to address specific regional programs to balance out regional inequalities which might emerge.

Generally, outcomes reinforce previous findings from the literature that demand linkages are important to explain regional wages disparities. The main and different result is related to trade shock impacts, which seem to have changed the strength of internal and external market. After both trade cost reduction, internal market appeared to have its explanation power weakened, while external strengthened.

Not only might trade policies influence firms' location, but also other government interventions may be able to affect firms' decisions. If a firm decides to expand or create a new plant, it might choose other locations for its production plant. Some of them may be credit constrained to implement a new project; therefore government support may help them pursue their strategy plans. If this government support helps them become more productive, then regions where the project was implemented might benefit from these supports. An increase in firms' productivity leads to growth; therefore firms and individuals are attracted to those growing regions.

Chapter 4 investigates whether government policies are able to improve firms' productivity. Evaluating government intervention on firms' performance and development is still in its infancy. Many tools to evaluate government policies have been developed by other branches of the economic literature, such as labour economics. Nevertheless, those methods have started to be implemented to evaluate government policies in firms' performance only recently. It is therefore a branch of the literature which has many venues to be explored. Among all of those, one vein seems to be extremely important: productivity. In the growth literature, an increase in firms' productivity appears to be one of the most important issues. Krugman (1992) points out for example that productivity growth is the crucial issue in the long run, including by comparing it with many others, such as lack in

infrastructure or technology gap. Examining whether government policies might help to improve firms' performance is then relevant for understanding how those should be designed.

In the developing world, firms are generally credit constrained. Even when they are not completely credit constrained, they are at least more restricted to finance resources than in the developed world. Evidences from the literature show that this is the case. Moreover, Brazil is not an exception to this rule. One way the Brazilian government has to relax these credit constraints is to provide long-term loans through its development bank (*Banco Nacional de Desenvolvimento Economico e Social* [*BNDES*]).

Among the loans available to entrepreneurs, two are designed specifically for the creation of new plants and/or the enlargement of existing ones: FINEM and Automatic BNDES. The main difference between these two is whether the loan is directly lent by BNDES (as for the former) or by utilizing retail banks (as for the latter). These two loans may impact however firms' productivity since they aim to improve their performance either by expanding existent plants' capacity or building up a new and modern plant. Chapter 4 investigates whether this government support was able to improve firms' productivity in the manufacturing sector.

One of BNDES's statuary goals is regional development. Poorest regions in Brazil receive therefore relatively more support than those developed. It is important to examine whether any difference exists between projects implemented in poor or rich regions. Another interesting issue is project size, since small and large projects may have distinguishable outcomes. Having this in mind, Chapter 4 explores these two channels: project size and location. In other words, the analysis involves whether the government policy is more successful in a more or less developed region and/or whether small or large projects are more associated with these bank loans. It is important to note that any

combination of these two channels, small projects in a poor region for example, is also investigated. These government loans are therefore evaluated not only through all types of projects developed in the Brazilian economy but also which size and/or region is mostly associated with them. In order to do that, firm-level data from the manufacturing sector from a diverse range of sources are used for this purpose from 1995 until 2003.

Different methods are implemented, from the most naive model to the more sophisticated. At first, a simple comparison between granted and non-granted firms is presented. As figures show, supported firms are not only more productive before borrowing from BNDES but also improve their performance after signing the contract. On average, they grew annually 8% more than non-granted firms. In order to evaluate whether this performance is associated with government financial support, econometric specifications are used for this purpose. A naive model is then performed to examine whether firms' performance is associated to BNDES loans. Overall, firms' performance is associated to BNDES loans in any scenario from projects' size to their location.

Firms may have unobserved attributes which might explain why supported firms perform better than others. By distinguishing between granted and non-granted firms' unseen characteristics, another method is implemented. Outcomes show that only small projects or in rich areas or both jointly (small projects in rich areas) are positively associated to government interventions. However, firms' unobserved heterogeneity may be greater than just a binary distinction between granted and non-granted. Panel data fixed-effects approach tackles this shortcoming by allowing unseen characteristics to vary across all firms throughout the whole period. Negative association appears by doing so. Without any covariate, which means only controlling for unobservable heterogeneity between firms, results show negative correlation in small projects and the same project in poor areas. On the other hand, large projects overall Brazil and in rich areas appears to be positively

correlated to BNDES financial support without controlling for all observable attributes. After controlling for all observable attributes, the only statistical significance result is a negative association between BNDES loans and firms' performance in poor areas.

Instead of checking whether the level of productivity is associated to government financial support, it is important to examine the association with their growth. First differentiating the data, apart from BNDES dummy, allows us to test this relationship. It is not possible to reject the hypothesis that no association exists between government financial support and firms' productivity growth by using the data available.

All those methods described so far assume a linear relationship between firms' performance and government support. One way to relax this assumption is by matching granted firms with non-granted ones. Aside from that, approaches used up to now are able only to infer association, not effect. Difference-in-differences approach is extensively used in the literature to fill up this gap. A combination of matching and dif-in-dif is then implemented in order to tackle these drawbacks. Results show they fail to reject the hypothesis of BNDES loans effect on firms' productivity in any venue, either by projects' size, or their location. Two possible conclusions emerge from this analysis. One is that BNDES loans are not able to improve firms' productivity over the period analyzed. Or simply that any conclusion cannot be achieved from government interventions on firms' performance, since those methods only fail to find any effect, neither positive nor negative. And whenever results fail to reject a null hypothesis, definite conclusions are not able to be made since uncertainty prevails. As one way that science progresses is by refuting a hypothesis made, in this case no effect; therefore failing to reject the hypothesis only provides us evidence that further work should be pursued in order to examine this issue.

Aside from this introduction and the other three chapters just mentioned, Chapter 5 concludes.

Chapter 2

Amenities and Migration Flows

Migration is as old as human beings. Moreover, reasons for changing dwellings have not changed dramatically over the years, since the majority of moves relate to individual decisions about where to maximize utility by looking at economic, social and environmental aspects. The majority of these aspects has been studied in previous works, as shown in the survey of Greenwood (1997) and Lucas (1997). Nevertheless, some areas are not fully covered by the existent publications. A vein which has not been completely explored yet is whether amenities are linked to migration. This chapter contributes to the literature by investigating the relationship between migration flows and amenities/disamenities, especially violence.

Amenities are divided into two classifications: natural and social. Natural amenities depend on the geographic characteristics of a particular region, such as natural beauty, temperature and location. Social amenities result from interactions among individuals in a particular location. Natural amenities, in similar fashion to social amenities, vary across regions. However, the latter vary across time, while the former do not.[3] This is an important issue in understanding migration, since as social amenities change one region may change its status from that of a desirable place to that of undesirable one. Examples of social amenities/disamenities include *inter alia* varieties of entertainment (such as cinemas, theatres and shows, for example), pollution, congestion and violence.

[3] Actually, natural amenities might vary across time; since one of the world's recent challenges is climate change. Social amenities are however much more time-variant than natural amenities. Therefore, assuming that natural amenities do not vary across time seems reasonable, as those may be considered static while comparing to social amenities.

Amenities may present an additional feature: neighbouring effects. Although natural amenities might show neighbouring effects, social amenities present relatively much more since these result from social interactions. Therefore, neighbouring effects of social amenities might be interesting to address as they are not restricted to the region where they are located. Social amenities may hence have an impact not only where they are located, but also in surroundings areas.

Violence is a social disamenity, which impacts on urban issues, as pointed out in Habitat (2007). According to this publication, people are even more affected by this disamenity in Less Developed Countries (hereafter referred to as LDC), including psychologically, socially and economically. Brazil is more than just a typical example of these effects because crime rates have achieved outstanding records. According to the UN Office of Drugs and Crime (UNODC), homicides rates in Brazil are one of the highest in the world, as it states: "Levels of small arms violence in countries at 'peace' can be as high, or even higher, than levels in war zones. For example, total gun deaths in the city of Rio de Janeiro between 1997 and 2000 exceeded conflict deaths in war zones such as Afghanistan, Sierra Leone and Uganda during the same period." in IANSA (2007), page 4.

Crime rates are not spread evenly over the country. As Carvalho, Cerqueira and Lobao (2005) and Waiselfisz (2007) have shown Brazilian violence is spatially concentrated in metropolitan areas. In 2007, the Brazilian federal government launched a national plan to tackle this issue, called PRONASCI,[4] which aimed to reduce crime rates in the 11 most violent metropolitan areas. Additionally to the criminality issue, Andrade, Santos and Serra (2000) findings show that the population in Brazilian medium size cities increased their participation in urban population from 9% in 1970 to 14% in 1996. As they are demographically less dense places, perhaps some migration towards these medium-

[4] PRONASCI is an acronym for *Programa Nacional de Seguranca Publica com Cidadania* which means National Program of Public Security with Citizenship.

sized cities might be associated with violence. Violence might thus have influenced people to migrate towards places with higher security.

In order to investigate how migration decisions are associated to amenities/disamenities, especially violence, urban to urban migration flows between Brazilian cities[5] are used from 1985 to 2000. These linkages are tested not only locally but also their neighbouring effects. Results suggest that they seem to be associated. Disamenities tend to push migrants towards other cities, while amenities seem to pull them away. As it regards violence, local crime rates are negatively associated to migration flows, but neighbouring does also present similar pattern. Controlling for origin and destination characteristics, crime rates at origin seem to be more associated to migration flows than their value at destination.

The rest of the chapter is structured as follows. The next section presents some stylized facts on migration and violence in Brazil. Section 2.2 presents some theoretical models to explain migration patterns and why crime is unevenly geographically distributed. The empirical strategy is detailed in Section 2.3. Section 2.4 presents a data description, followed by results in Section 2.5. Finally, the last section concludes.

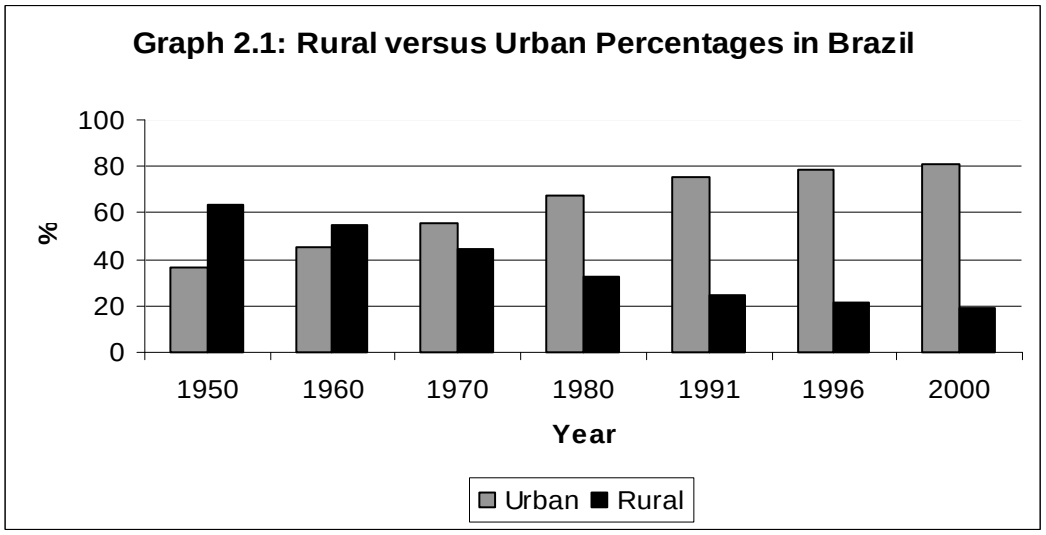
2.1. Some Facts about Migration and Violence in Brazil

2.1.a Migration in LDC and in Brazil.

Migration differs not only between rich and poor countries, but also across time. Most of the literature regarding migration in LDC investigates rural-urban migration, as summarized in Lucas (1997). Even recently, papers addressing this type of migration either theoretically (Chaudhuri (2000) and Basu (2000), for example) or empirically (Bhattacharrya (2002) and Saracoglu and Roe (2004)) have been common. The trend towards a more urban society is however also something common in LDC. In LDC

[5] Overall, there are 3,659 regional units.

countries, urban population was 29.6 % in 1980, by 2005 the percentage increased to 42.7% according to United Nations (2007). In South America, urban population was even more important, where 81.8% were living in cities in 2005 compared to 68.3% in 1980. Brazil is not an exception to these figures, as Graph 2.1 shows:



Source: Brazilian Statistical Institute (IBGE)

Brazil has experienced a substantial change in people's residence status in the last 50 years. In the 50s, two out of three Brazilians were living in rural areas.[6] However, the urban population became larger than the rural in the 70s. It has steadily increased since and more than 80% of Brazilians live in cities nowadays. This indicates a strong urbanization process has occurred in this country. Nevertheless, it is not possible to infer whether migration from rural to urban contributed massively to this process.

There is extensive literature about migration in Brazil and only papers which are directly related to the present study are mentioned here. As mentioned earlier, medium-sized cities are becoming more representative in the Brazilian urban population. It is not clear however which demographic channel (migration, birth and/or death rates) explains this pattern. Contributing to this issue, Jannuzzi and Jannuzzi (2002) analyse migration,

[6] The definition of urban population according to IBGE encompasses any person living in dwellings located in municipalities, villas or even isolated agglomerations of dwellings regardless their size, demographic density or any other criterion. Rural dwellings are those not situated in these urban areas.

urban growth and real estate attractiveness of Sao Paulo, the economic centre. They find that Sao Paulo lost inhabitants due to the migration process which occurred from the 1980s onwards, as evidenced in Martine (1994) and Baeninger (2000). One of Jannuzzis's explanations is that housing prices are becoming substantially higher, especially for low income families. They evaluate the association between house prices (represented by rent) and different types of measures, such as time spent on travelling to work, distance to downtown, population density, *inter alia*. Even though all of these measures are important to understand migration patterns, they use them only in a simple linear regression model which leads to bias estimated parameters, due to omitted variables. Additionally, they do not encompass violence. They conclude that people migrated to cities around Sao Paulo in order to avoid high house prices. Their paper provides one indication that migration towards medium size cities can be related to people leaving metropolitan cities.

The closest paper to this one is Vasconcellos and Rangel (2005). They analyse Brazilian migration patterns using the last two Census data as does this chapter.[7] They provide a detailed descriptive analysis of the migration process in Brazil during this period using city level data. Comparing both periods, the authors conclude that migration flows have remained the same within the period analyzed. They find moreover not only that 10% of Brazilians over five years of age have migrated from 1986 to 1991 and from 1995 to 2000, but also that people from the southeast and the northeast were the majority of them (65% in 1991 and 71% in 2000). The population of the southeast has the greater positive net result for migration flows in contrast with the population of the northeast which revealed the largest negative net result for migration flows. The average distance is slightly greater than 400 km, but more than 60% of migrants change their residence within 250 km, of which half (around 30% of the total) do so within 50 km. A reasonable part occurs from

[7] Section 2.4 explains the use of Census data instead of other sources.

metropolitan cities towards their neighbours, which can be interpreted as individuals looking for better house pricing and/or better amenities. Another finding is that the level of intra-state migration (between cities within the same state) is significantly higher than the level of inter-state movements (across cities in the same state). All of these together indicate that distance matters. Considering all possible migration flows, around 1% is non-zero. One issue not covered by these two papers just mentioned is the identification between rural and urban as origins and destinations, which is considered in this current chapter.

2.1.b Urban versus Rural

It is common sense that rural-urban migration is the standard pattern in LDC as Lucas (1997) states, but developing countries may present completely different migration flows. India, for example, is an exception to this rule. Skeldon (1986) shows that 57% of migration in India was rural-rural according to their 1981 Census. Focusing on rural-urban migration, Nelson (1976) argues Latin America differs from Africa/Asia by migration period status. In the latter, it is basically transitory and in the former, permanent. This explains partially why Latin America is much more urbanized than other parts in LDC. Considering this, it is possible to expect that rural-urban migration might be less important for migration flows in Latin America, especially in Brazil which presents higher urban rates. Table 2.1 shows Brazilian migration flows by distinguishing between rural and urban migration as origin and/or destination.

Table 2.1: Numbers and Percentage of Migration in Brazil

	1991		2000	
	Migrants (in millions)	%	Migrants (in millions)	%
Urban to Urban	8.0	62%	9.9	72%
Urban to Rural	1.1	8%	1.1	8%
Rural to Urban	2.3	18%	1.7	13%
Rural to Rural	1.5	11%	0.9	7%
Total	12.9	100%	13.7	100%

Source: Censuses of 1991 and 2000, elaborated by IBGE.

The table presents how migration from rural to urban still plays an important role, but it is more than evident that urban-urban is much more relevant than any other. While migration from rural to urban has decreased not only in volume (from 2.3 million to 1.7 million), but also in share (from 18% to 13%), migration between urban areas has increased in both: volume and share. This issue is not investigated by Jannuzzi and Jannuzzi (2002) and it presents a completely different pattern between the two censuses analyzed, since urban-urban migrations become even more important than rural-urban. In total, around 13 million people change their residence every five years according to the two censuses, representing around 10% of the Brazilian population, as pointed out by Vasconcellos and Rangel (2005). Indeed, a volume of the entire Brazilian population changed their address to another city every half century after considering these figures. People move therefore 1.4 times during their lifetime if Brazilian life expectancy is considered.[8]

2.1.c Violence Issues

As it regards violence, the urban population is increasingly being affected by this disamenity, as reported in UN-Habitat (2007). According to this publication, urban violence is at the top of the agenda for people living in cities, especially in developing countries where a huge proportion of them live in poor areas with lack of infrastructure and labour opportunities. Another example of how much violence affects society is given by

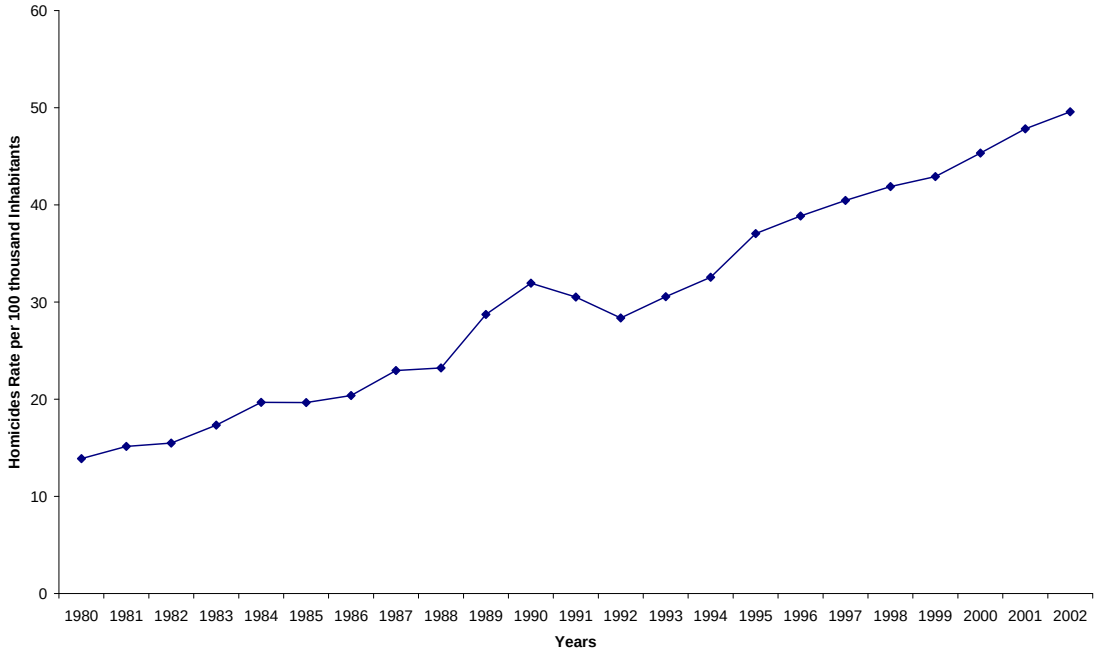
[8] According to IBGE, Brazilian life expectancy was 70 years in 2000.

Soares (2006). He estimated the cost of violence in the welfare of 73 countries. His findings state that one year of life expectancy lost to violence is associated with a yearly social cost of 3.8% of GDP.

Across Latin American countries, and especially in Brazil, violence also represents an important issue in urban areas. Carvalho, Cerqueira, Lobao and Rodrigues (2007) states that the cost of violence has reached 5.09% of Brazilian GDP in 2004. Aside from economic issues, the UN-Habitat (2007) also mentions how the fear of violence affects people's life, since they state that even when crime rates reduce, the sense of insecurity may remain high, hence lessening people's quality of life. Social life can also be affected by violence. For example, more than 60% of Brazilians feel unsafe walking home at night, according to Nuttall, Eversley and Rudder (2002). Summing up, crime imposes economic, social and psychological negative effects in the Brazilian Society.

In order to have an idea of how violence has developed in Brazil, Graph 2.2 presents homicides rates per 100 thousand inhabitants from 1980 until 2002.

Graph 2.2: Brazilian Homicide Rates from 1980 until 2002



Source: SIM-DATASUS, Brazilian Health Ministry

As shown, there is an upward tendency of this crime over the whole period. It started at a rate of roughly 14 deaths per 100 thousand inhabitants and reached nearly 50 per 100 thousand inhabitants which represent an annual increase of 6%. This demonstrates how this issue has become so prominent for the Brazilian society over the last couple of decades.

The distribution of crime varies substantially across municipalities where some present reasonably low crime rates while others resemble armed conflict areas. In order to evaluate how violence is distributed across these regional units over the period, violence data is available; Table 2.2 reports statistics of the five-year average homicide rates from the 3,659 cities under study.

Table 2.2: Descriptive Analysis of Violence in Brazil

Homicides per 100 thousand inhabitants	1986-1991	1995-2000
Average	8.6	9.5
Median	5.6	6.1
Standard Deviation	10.2	11.6
Maximum	116	140
Minimum in Cities with Crime Record	0.24	0.34
Number of Cities with no Crime Record	624	524

Source: SIM-DATASUS, Brazilian Health Ministry

First, cities present on average lower crime rates compared to national figures, since homicides rates across cities are below 10 homicides per 100 thousand inhabitants while at the national level they are 26 from 1986-91 and 41 from 1995-00. This may represent higher crime rates in more populous regions. The upward tendency shown in Graph 2.2 is also noticed nevertheless at Table 2.2 by any descriptive statistics considered.

The average across municipalities has increased by one more death per year corroborated by an increase in the median as well. However, the increase has not been homogeneous across cities, since the standard deviation increased which means more heterogeneity across cities. Also the number of cities with no statistical record of homicide rates has reduced after a decade, as shown by the decrease in the number of cities with no crime record.

The minimum and maximum values observed also provide an indication of crime rates rising over the years in each locality. The maximum five-year average registered is from a city belonging to the Sao Paulo metropolitan region which may indicate that big metropolitan areas are in fact the most dangerous.[9] More generally, Carvalho et al. (2005) and Waiselfisz (2007) provide evidences that areas with high crime rates include the most important metropolitan areas, such as Rio de Janeiro and Sao Paulo. It is thus evident that the more populous regions are among those with the higher crime rates.

Waiselfisz (2007) explores how geographically violent rates are distributed over Brazil. One of his findings show that the top 10% more violent cities concentrates more than 70% of all homicides rates in 2004. Moreover, these municipalities are generally big cities as those cities represent 42% of the Brazilian population. Taking into consideration the average homicide rate in 2004, those 10% more violent cities are four times more violent than the national figure.[10]

Although Waiselfisz (2007) shows graphically that violence is spatially concentrated, they do not test it statistically. Carvalho et al. (2005) not only present some maps to show this pattern, but they also investigate whether Brazilian homicides rates across municipalities are spatially correlated. Their findings back up the map evidence. This violence spatial correlation allows us to consider that violence might have neighbouring effects.[11]

Violence seems to impact not only where it occurs, but also in neighbouring areas. One example of how violence might affect economic measures, locally and/or in surrounding areas, is given by Gibbons (2004). Gibbons tries to investigate whether

[9] Sao Paulo metropolitan region is the most populated region in Brazil, and Diadema is the city where crime rates are beyond 100 homicides per 100 thousand inhabitants at both periods.
[10] This explains why national figures present a higher increase in homicide rates, as shown in Graph 2.2, than the average homicide rates across municipalities, as Table 2.2 presents.
[11] In the Section 2.2.b, this assumption of crime spatial correlation is theoretically modelled by many papers having a diverse range of arguments to explain this pattern.

housing prices may be affected by violence. Using a neighbourhood data set, his results suggest that violence, locally and in surroundings areas, affects housing prices in London. Violence seems therefore to have its spatial effects on economic issues.

With regards to migration, violence might have a spatial effect due to some reasons. For example, even living in a city with low levels of violence, if neighbouring areas show high criminality records then inhabitants of this safer area might fear that sometime in the near future violence might increase in their locality. Therefore, spatial effects of violence may influence people's decision to migrate. Another feature is that people may work and live in different cities as happens in some metropolitan areas in Brazil. Although people may live in relatively peaceful areas, they may work in areas where criminality levels are higher. Therefore, facing more violence during the work day might affect people's decision to move to another area. If this happens, neighbouring effects of violence are influencing people's decision as to where to live.

Considering urban-urban is the most common type of migration for the Brazilian case, it is possible to analyze whether these movements are associated to amenities/disamenities, especially violence. The preliminary descriptive statistics presented in this section provide some indication this might be the case.

2.2. Theoretical Background

2.2.a Migration Theory

A theoretical framework may help us to understand how migration and amenities/disamenities, including violence, are correlated. New Economic Geography (NEG, henceforth) is a branch of the literature which investigates how migration influences

population distribution across regions.[12] Generally, there are two types of workers: migrants; and non-migrants. Mobile workers decide where to live considering regions' attributes, therefore determining each region's size.[13]

One seminal paper of this literature is Krugman (1991). In his work, migration of workers determines how economic activity is distributed across regions. In his model, there are two types of workers: manufacturing workers, who are mobile across regions; and farmers, who are immobile.[14] Only manufacturing products faces iceberg transport costs to ship their goods to consumers. Since farmers are immobile, there is an extra assumption stating that they are equally divided into these two regions. When transport costs are low in this model, there are three equilibriums, but only catastrophic results (agglomerating in only one region) are stable, the symmetric one is unstable. Any disturbance in the symmetric equilibrium will lead therefore workers to migrate towards one region. When transport costs are high, symmetric equilibrium is stable and catastrophic equilibriums are unstable. Five equilibriums appear when economy faces intermediate transport costs, where three (two catastrophic and symmetric) are stable and two are unstable. It is hence possible that migration will be towards agglomeration in one single economy or dividing equally. Real wages are the main factor determining where mobile workers decide to live. If one region has real wages higher than the other, workers tend to migrate hence from the lower salary region to the higher one. The rate of migration is determined moreover by the difference in real wages expressed in the equation below:

$$\frac{\dot{\lambda}_r}{\lambda_r} = \gamma\left(w_r - \overline{w}\right) \tag{2.1}$$

[12] Although migration equation is ad hoc in NEG models as pointed out by Baldwin, Forslid, Martin, Ottaviano and Robert Nicoud (2003), they have the advantage in explaining migration by considering how origin and/or destination are affected.

[13] Employees live and work at the same location.

[14] Even though this appears to be a strong assumption, it seems that this might be a reasonable approximation if we look into the Brazilian data, since most of migration occurs between cities, where manufacturing workers live and work, and not between, from or towards rural places.

Where λ_r is the proportion of manufacturing workers in region r; w_r is the real wage in region r; $\overline{w}$ is the average wage across regions; γ is a parameter which address the question whether wages interfere in migration

Equation (2.1) represents location decisions taken by individuals in terms of real wages at two locations, since they determine the utility level of workers. Depending on the level of transport cost, full agglomeration or an even distribution of migrants may occur. If transport costs are low, then full agglomeration equilibrium not only occurs, but it is also stable, while symmetric distribution is unstable. As transport costs rise, the greater the chance is of having an even distribution of migrants and lower is the possibility of full agglomeration on one side. These results lie on a crucial assumption of the model: homogeneous consumers. This assumption leads to the conclusion that catastrophic agglomeration occurs when trade costs reduce, since workers migrate towards regions with higher wages.

Contributing to this literature, some NEG models assume taste heterogeneity for consumers which operate as a dispersion force preventing full agglomeration. People with heterogeneous taste prefer to live in distinct areas by valuing regions' attributes differently. Murata (2003) is one example of this literature. In his model, he assumes not only taste heterogeneity but also eliminates the hypothesis of immobile factors. His conclusions show stable dispersion distribution of economic activity, and therefore workers, is reached when transport cost decline in all different scenarios of taste heterogeneity.

Tabuchi and Thisse (2002) is another example of using taste heterogeneity as a centrifugal force in NEG models. Even when trade costs reduce towards zero, catastrophic agglomeration does not occur due to this centrifugal force (taste heterogeneity), as found in Murata (2003). An additional feature investigated by the authors is whether this catastrophic agglomeration occurs when regions differ. In other words, when one place has

more natural amenities than the other, this area will attract migrants and the other repels them. Their results remain practically similar to previous assumptions and places with more amenities generally present a larger population. Workers may decide however to locate in a region with a low level of amenities when: (i) transport costs are at intermediate levels; and (ii) when differentials in taste heterogeneity and amenities are not substantially high to be distinguished by workers.

This paper focuses only on natural amenities, such as temperature and scenic beauty. Social amenities are not treated since they vary according to population size. They acknowledge the absence of social amenities as a limitation to their model. Authors' justification for not including social amenities into the model is that they depend on the number of residents due to their increasing returns to scale production. Other social amenities result from agglomeration, especially in demographically dense areas, such as violence and congestion. How important natural and social amenities for an individual decision to move is a challenging question, as they state. The answer for this question varies moreover according to the development stage of a particular region which is being investigated.

Therefore, this paper investigates whether migration decision are associated to those social amenities, which are not easy to model. Nevertheless, results should bear out similar outcomes suggested by natural amenities, which are positive attracting and negative expelling inhabitants.

2.2.b Violence

Looking at violence, Glaeser, Sacerdote and Scheinkman (1996) point out crime rates vary across space and time because social interactions create reasonable covariance across individuals. Places where criminals have more social interaction therefore tend to have high violence rates. The main intuition behind their argument is that social

interactions seem to create a sense of invulnerability and a willingness to violate social norms and take risks, as long as one is in the company of like-minded comrades. This idea backs up some previous findings. According to Reiss (1980), two-thirds of all crimes are committed by offenders acting alone, but the exact same percentage of criminals commits their crimes with a peer. In order to investigate this social interaction related to crime, Glaeser et al. (1996) create a theoretical model in which social interactions occur at a local level and agents' decisions are influenced therefore mainly by their neighbours' actions. Their model contains not only local interactions, such as informational spillovers and/or behavioural influences, but also global interactions, as well as labour market conditions and police expenditures. Their theoretical findings are corroborated by performing an empirical test using US data in which the main result states that a reasonable amount of social interaction exists in criminal behaviour even after controlling for a diverse range of cities' characteristics. Across cities they found that regions with high crime rates are those where family units are not complete, such as single-parent families.

Sah (1991) also address this issue by establishing a benchmark model which explains the following: in places where police is limited to catch criminals, an individual's choice to become a criminal lowers the probability of any other individual to be arrested. Therefore, where more crime occurs the probability of being arrested reduces. His model ends up having two (but possibly more) equilibriums, either places with high crime rates and low probability of being caught by police, or high probability of being arrested and low levels of criminality. Suggesting a similar and alternative approach, Murphy, Shleifer and Vishny (1993) also find multiple equilibriums. Their idea is high levels of criminal behaviour crowd out legal activities at any location. If the number of criminals increases, then returns of non criminals reduce since legal activities are stolen by criminals.

These are papers which address why crime is not evenly distributed in geographical terms, but none of them link migration with crime. These findings on violence suggest that this social amenity affects not only where it occurs, but also in its surrounding areas. If violence has some spatial interactions, then it might affect migration flows.

Regarding how violence might be affecting people's decision on where to live, there has been some work on how city growth and/or size have been impacted by crime rates. One seminal paper is Cullen and Levitt (1999), where they investigate how population size has responded to violence in the USA. Their findings suggest that each additional reported crime is associated with a roughly one-person decline in city population. Although migration is the main part in explaining city size, it is not the only part as birth and death rates play also their role at city size. In order to overcome this shortcoming, the authors also estimate how violence impacts net migration and results remained the same. However, their main drawback is the fact that they do not have information on origin and destination of movers. Therefore, their paper cannot address directly how origin and destination are affected by crime rates. Another example of this literature is Da Matta, Deichmann, Henderson, Lall and Wang (2007) in which they estimate what are the determinants of city growth in Brazil. Among those determinants, violence is one variable investigated. Their outcomes suggest that crime rates affect negatively city growth, showing that places experiencing higher violence tend to grow at a lower rate. Nevertheless, it is important to mention that they do not address specifically migration on their paper, since they use only city size as the previous paper mentioned in this paragraph.

It is therefore important to notice that this interaction between crime rates and migration needs to be further investigated. Aside from this, Brazil seems to be an interesting case to investigate this phenomenon because medium-sized cities with lower

levels of criminality are becoming more relevant. Moreover, violence rates reached outstanding rates.

2.3. Empirical Strategy

The socio-economic phenomenon proposed in this chapter is to what extent amenities are correlated to migration flows. In order to capture this phenomenon, it is advisable to control for many other socio-economic variables aside from amenities. Before specifying the econometric equation, discussion on which variables might be important remains relevant.

Migration is a worldwide phenomenon, occurring in the developed or developing world, and individuals from both classifications might have similar incentives to migrate, but there are also relevant differences. For example, migration in developing countries is not fully explained by formal job opportunities since the informal sector plays an important role. In developed countries, access to affordable housing might act more as a substantial barrier than in poor countries where shanty towns are created by low-cost housing projects. Variables used to understand migration should therefore be similar regardless of which particular region is investigated, yet some others must be included in order to adapt to the reality of each region.

The list of relevant variables to explain migration flows is described as follows: economic variables are discussed first, followed by amenities, social issues, infrastructure and finally particularities of the investigated region. First, it is important to have a variable which might represent the cost to migrate. In the migration literature, distance is used for this purpose since it represents not only financial costs but also time spent to actually migrate.

Another common economic issue for migration is wage disparities. Regional differences in wages create incentives for migration, since regions offering higher wages attract workers. Regions offering higher nominal wages might however have higher living cost, because of housing, for example. Therefore, the main variable should be real wages, rather than nominal wages. Skills necessary to work might differ between distinctive sectors. For example, skills needed to work in the service sector may be different than in the manufacturing sector. Disaggregating salaries among sectors may then be relevant. However, it is important to control for the importance of each sector in every municipality since higher wages in one locality may be biased by the existence of a single firm which hires too specific a skilled worker.

In order to control for how many job opportunities arise in each region, the unemployment rate may be considered to tackle this issue by signalizing which area may be creating fewer chances to work. The development of a particular location is also relevant to be taken into account since faster-growing regions tend to create more job opportunities. Subsequently, the growth of each region might also be relevant to include. According to the NEG literature, market potential is relevant to attracting migrants due not only to the existence of a greater labour market but also because it may capture the effect of the number of varieties of goods and services in each location, since cities with higher market potential provide more of them to their consumers. Not only internal market might be relevant for migration decision but also external market. Some regions may be affected by an external shock, such as a boom in exports, and then using a measure to isolate the foreign shock effect may be relevant.

Aside from economic reasons, amenities/disamenities might be also an important factor to attract or to repel migrants, but some empirical work provides some results which bear out to be careful when incorporating them since they might be correlated to other

economic variables, such as wages and/or housing prices. One part of the literature argues that wages may be able to incorporate many regions' attributes, which is the assumption underlined in the NEG models as they are the main factor to explain migration decisions. At first, people may find this assumption extremely strong, but it seems to be a good simplification of reality corroborated by some empirical studies, such as Rosen (1979). In his paper, Rosen addresses how amenities/disamenites for example might explain regional disparities in wages. He assumes that local attributes vary across cities, but not across time: climate is a good example that satisfies this assumption. He points out two types of places: desirable and undesirable. The desirable places generally have higher housing prices and lower wages relative to the undesirable places.

His model describes acceptance-wage gradient by a linear combination of the real productivity effect and pure taste effect. His conclusion is that cities with more amenities should consequently have lower wages in relation to places with fewer amenities. He also develops a bid-rent function which uses the same attributes (real productivity effect and pure taste effect) and his conclusion is that cities with higher amenities will have higher rents.

One of his assumptions is that people have different tastes; hence, they decide to live across cities according to their tastes. In order to evaluate his theoretical findings, Rosen performs an empirical test to evaluate how wages in most metropolitan areas could be explained by some city attributes by using individual-level data. First, he finds no difference using nominal or real wages. It should be pointed out however that wages are from metropolitan areas which do not have too many discrepancies of living cost between them. Regarding city attributes he uses a diverse range of measures from pollution, climate, crowding, and market conditions. For crime data he utilizes only the total crime rates from the FBI. Even though the author admits to having details of each criminal activity, he does

not use this extra information. He uses only an average measure of all crimes and does not report what the types are and how he combined them. He finds that most of the attributes are significant to explain wages, such as pollution, climate, and market conditions. Crowding and crime do not provide robust results in affecting wages.

Some other papers address these similar issues and also include crime rates to be captured by wages, but different results emerge. While Berger, Blomquist and Hoehn (1987) find that an increase in violence reduces wages, Clark and Cosgrove (1991) and Blackaby and Murphy (1991) show the opposite. Hence, one conclusion which arises from these findings is that natural amenities might really be captured by wages, while social amenities not, especially violence.

If regional differences in wages may be explained by distinct city attributes, these findings on violence suggest that certain social amenities might not be properly reflected in salary disparities, since it is not clear how wages and violence, for example, are associated. Rosen explains his violence result by stating that the crime rate depends on the precise location within the area where it was committed. He overlooks however other issues rather important to explain his findings. First, the two city attributes presenting unexpected results, crime and crowding, are not fixed over time and change according to population size. Wages may adjust to these characteristics, but there might be a period for this adjustment. In other words, it is not instantaneous. Salaries can only adjust to the last crime rate reported (time $t-1$) at time t, but this disamenity may have changed by time t when the adjustment occurs. Wages may not therefore capture perfectly these time varying amenities/disamenites, such as crime. Labour market frictions moreover may prevent this adjustment from occurring. Even if salaries could adjust instantaneously to these measures, crime rates might not represent people's fears. It was stated previously that even when crime rates reduce, perception of safeness might not change. People's feelings on security

are therefore even more difficult to be captured by wages because it is not measurable. Furthering this argument, people react differently to these social amenities since some residents find a place more pleasant if they are able to see people walking on the streets constantly, while others prefer silent places. How people value these characteristics may differ quite substantially. This point is not fully covered by any of these papers mentioned beforehand, since none of them consider that some amenities (especially the social ones) vary across time. As a consequence, violence might play a role in migration flows.

Summing up, any natural amenity is already being captured by wages; therefore it is not needed to control for them. On the other hand, social amenities should be incorporated as explanatory variables.

Apart from economic reasons and amenities/disamenities, social issues may also be important. Social network is also a key variable for migrants, because it reduces the migration cost as those networks are able to inform newcomers about labour and housing markets, for example. The literature of migration uses the share of urban population in each locality as a proxy, since this measure captures social interactions among people. In other words, the more people living in one area the higher the probability is of finding a relative and/or a friend who might help to smooth migration costs. This measure may also represent however other social amenity/disamenity occurred in a particular region. The bigger a place is, for example, the more services are provided, such as restaurants, bars, cinemas, theatres among others. It might also represent nevertheless disamenities, such as crime, congestion, and pollution. In order to isolate some of these amenities' interferences, some measures of them are required. Regarding the main focus variable of this chapter, violence - although the literature and empirical works show bigger metropolises tend to have higher crime rates - it is important to isolate violence to examine whether it is correlated to migration flows or

not. Another social issue related to migration is gender. Overall, males tend to migrate more than females, which enables a creation of gender variable to capture this effect.

As people move to places to maximize their utility, measures of infrastructure may also be relevant to include. One social infrastructure which might be relevant is health access, since places with good access to health care might be important for migrants' decision, especially if they care substantially about violence. Another relevant social infrastructure is education as some migration decisions are made in order to get access to higher degrees, as happens with people moving to undertake a university degree. Measures of physical urban infrastructure may also be convenient to include as controls. It is thus important to use that infrastructure information in order to evaluate how migration is affected by them.

Regarding some particularities about this specific country, two emerge. One common issue related to migration decision in Brazil is job opportunities in the government sector. As those are generally in the state or federal capital, a measure to capture this type of migration seems to be necessary. As said previously, informal sector plays an important role in developing countries and Brazil is not an exception to this rule. It should be relevant to control for this country's particularity.

2.3.a The Econometric Model

Generally, migration is empirically investigated by using the gravity equation where information of origin and destination might be used to explain it. The econometric specification changes however depending on the hypothesis made. If the hypothesis assumes people have decided to migrate from where they live and the only uncertainty is where to migrate, then the econometric specification should have information only from destination as explanatory variables, as shown in Cushing (1989). Poncet (2006) and Crozet

(2004) are examples of this approach. Both papers derive their econometric specification based on this hypothesis where individuals compare cities' characteristics of all possible destinations. As a consequence, their equation to be estimated is a restricted model of a gravity-type (benchmark) adapted to the NEG framework. Their model was used as a starting point to construct the econometric specification of this chapter and the version used is presented in equation 2.2 using the main variables investigated.

$$\ln shm_{ij,t} = \beta_0 + \beta_1 \ln wage\_manuf_{j,t-1} + \beta_2 \ln wage\_serv_{j,t-1} + \beta_3 \ln unempl_{j,t-1} + \beta_4 \ln dist_{ij} + \beta_5 \ln shpop_{j,t-1} + \beta_6 \ln MP_{j,t-1} + \beta_7 \ln viol_{j,t-1} + \beta_8 \ln PX_{j,t-1} + \beta_9 \ln NX_{j,t-1} + \gamma Controls_{jt} + u_{ij,t} \quad (2.2)$$

where, is the share of migrants from origin *i* to destination *j*; $shm_{ij,t} = \frac{mig_{ij,t}}{\sum_{it} mig_{ij,t}}$

$wage\_manuf_{j,t-1}$ is the average real wage of manufacturing sector at destination *j* at *t-1*; $wage\_serv_{j,t-1}$ is the average real wage of service sector at destination *j* at *t-1*; $unempl_{j,t-1}$ is the unemployment rate at destination *j* at *t-1*; $dist_{ij}$ is the distance between *i* and *j* ; $shpop_{j,t-1}$ is the share of urban population from destination *j* in the Brazilian as a whole at *t-1*; $MP_{j,t-1}$ is the market potential at destination *j* at *t-1*; $viol_{j,t-1}$ is the violence measure at destination *j* at *t-1*; $PX_{j,t-1}$ is any social amenity attracting migrants (number of bars, restaurants and theatres, for example) at destination *j* at *t-1*; $NX_{j,t-1}$ is any social disamenity repelling migrants (such as congestion, and pollution) at destination *j* at *t-1*; $u_{ij,t}$ is the error term.

As discussed above, the migration literature states that distance is negatively correlated to migrations flows since the farther a region is, the more costly it is to move, not only in financial terms but also regarding time. According to the theory, real salaries and

market potential should be positively associated with migration. People will therefore move towards cities with higher real wages and bigger economies. Labour market and migration literature say the unemployment rate should be negatively correlated to migration flows since jobless regions do not attract migrants. Social amenity should attract migrants, while disamenity repel them. Finally, violence should be negatively associated with migration, because people might tend to move to places with more security. To sum up, parameters are expected having the following signs:

- $\beta_1, \beta_2, \beta_6, \beta_8$ greater than zero;
- $\beta_3, \beta_4, \beta_7, \beta_9$ negative.

The only parameter which does not have a straightforward expected sign is β_5. It will be greater than zero when amenities (such as social network, represented by relatives and friends) overcome disamenites (like pollution or other types of crime), otherwise a negative sign might appear.

As was pointed out by Plane (2004), there is a lack of use of spatial models in the migration literature which is a problem because migration flows between any pair of regions are also potentially influenced by neighbouring characteristics from origin and destination. And the disamenity analyzed here is not an exception to this rule. Section 1.2.b provides empirical evidence of violence spatial correlation in Brazil. Additionally, Section 2.2 presented theoretical models explaining that crime has its spatial interactions. Therefore, violence may affect people's well being not only locally but also at surrounding areas. Spatial issues of violence might be thus an important issue in this investigation.[15] Improvements in the econometric specification are introduced in order to include spatial

[15] The average distance between the contiguous neighbors across these localities is around 30 km and 95% of these distances are below 74 km. It is therefore possible for an individual to move residence across these cities without quitting his/her job, which enables the use of spatial neighboring effects.

effects of violence since the effects of this disamenity spread across surrounding areas. The following model is eventually proposed in equation 2.3.

(2.3)

$$\ln shm_{ij,t} = \beta_0 + \beta_1 \ln wage\_manuf_{j,t-1} + \beta_2 \ln wage\_serv_{j,t-1} + \beta_3 \ln unempl_{j,t-1} + \beta_4 \ln dist_{ij} + \beta_5 \ln shpop_{j,t-1} + \beta_6 \ln MP_{j,t-1} + \beta_7 \ln viol_{j,t-1} + + \beta_8 \ln PX_{j,t-1} + \beta_9 \ln NX_{j,t-1} + \beta_{10} W \ln viol_{j,t-1} + \gamma Controls_{jt} + u_{ij,t}$$

Where W is a spatial matrix to taking into account neighbouring effects on migrations flows. For spatial lag, two types of matrices are used: contiguity; and four nearest cities. The reason of using the contiguity matrix is fairly obvious, since places surrounded by violent neighbours may not be considered as a destination for migrants. One possible criticism is that not only contiguous municipalities should be considered since two cities separated by a bay or river, for example, might have much more interaction because a bridge connects them than another two sharing a border.[16] Therefore, using the four nearest cities matrix overcome this shortcoming.[17]

One might argue that neighbouring effects should be considered in other variables, yet other drawbacks emerge on doing so. First, some variables do not have any spatial interaction, such as population density, since it is not clear that places highly densely populated should affect their neighbours. For example, cities in the suburbs around a metropolis should not necessarily suffer from it. Second, the spatial lag of other measures might be highly correlated to some others. One example is the spatial lag of other positive amenities: entertainment measure. The spatial lag of entertainment measure is being captured by market potential. Including these spatial effects might only interfere in market potential results. Last, but not least, there is no certain agreement on using spatial correlation in empirical research, as noted by Wooldridge (2002) "For better or worse,

[16] One example is Rio de Janeiro which has many more connections to Niteroi (which is on the other side of Guanabara Bay) than to Itaguai, which is a contiguous neighbour.
[17] The mode from contiguous neighbours is four and this explains why four nearest other cities are considered instead of three or five, for example.

spatial correlation is often ignored in applied work because correcting the problem can be difficult" (page 6). Having those points in mind, it seems reasonable to assume that only violence might have spatial interactions as theoretical and empirical papers have already evidenced, as shown theoretically in Section 2.2 and evidenced empirically in Section 1.2.b.

Another feature is that origin information may be essential for migration analysis. If so, the unrestricted traditional gravity equation should also be estimated. In order to provide a robustness check, a traditional gravity equation is also estimated using equations 2.2 and 2.3. Indeed, some differences from the previous model emerge. One difference is that the dependent variable is not divided by the sum of destination. In other words, the migration measure should only be the migration between *i* and *j*, and not a ratio. The figure is divided however by the total migration flows for each period in order to avoid any time trend.[18] The new dependent variable is described in equation 2.4.

$$shm_{ji,t} = \frac{mig_{ji,t}}{\sum_{t} mig_{ji,t}} \qquad (2.4)$$

Another change lies in the explanatory variables which now need to have not only the destination values, but also the correspondent measure in the origin. The market potential has for example two values in the econometric specification: one for origin and another for destination, where the destination is described as a pulling factor and the origin as a pushing factor.

The focus variable of this chapter, violence, is endogenous because it is not clear whether violence impacts migration flows or the other way around. Another issue is that the violence measure used here does not represent all crimes; moreover, it is the fear of crime

[18] As shown in Section 2.1, urban-urban migration increased substantially from 1991 to 2000 either in nominal terms or in share. Therefore, it is important to control for this time trend. Including just a time dummy is an option in which results do not present different outcomes. For simplicity, this chapter show only results using the dependent variable divided by the total migration flows.

which leads people to other places, and not crime rates. These shortcomings lead to biased estimates of the parameter related to this measure. It should therefore be tackled by using instruments in an IV estimation. Valid instruments should be those which can only affect dependent variable through the endogenous variable. Therefore, the exclusion restriction implied by any instrumental variable regression is that, conditional on the controls included in the regression, instruments have no effect on migration flows, other than their effect through violence. One potential instrument may be some policy intervention which affects directly the independent but not the dependent variable, apart from through endogenous variable.

Investment on security is a potential instrument, since it may only affect migration flows via its effects on violence. State governments are responsible for practically most of the Brazilian police system. The percentage of security expenditures on each state budget may be hence used to instrument violence. Although this might represent a response to criminality increase, some precautions are taken trying to avoid it. One strategy was taking information lagged in time. Therefore, only the level of security expenditure 5 years before Censuses is used for this purpose. By taking it lagged in time, government expenditures on security are previous violence records and migration flows. Therefore, time lagged information of security expenditure could be correlated with the violence measure, which may have a direct effect on migration flows. One additional feature of this instrument is its regional dimension (state level) which is actually larger than the unit used in this study (city level). It makes this instrument even more exogenous, since cities within the same state are considered similar, when they are actually much more heterogeneous, including in violence terms. Another issue is related to the spatial dimension of violence, since increasing crime rates in one city might be explained by the reduction in a neighbouring city. A government intervention in a broader geographical scale is able to tackle this issue.

This empirical approach might capture the effect of government expenditure on migration flows, but working through other channels. Furthermore, other variables might be correlated with security expenditure and migration flows. In order to tackle those issues, additional instruments might be used. Inequalities measures might be potentials instruments, since regions with worst income distribution may increase social tension, and therefore, more crimes. However, it is not feasible to think that someone might tend to move towards a poor city with lower inequality index, just because they would be equal everyone. Inequalities measures, such as Theil and Gini, at municipality level as well as at state level are hence used to instrument the violence measure, since they might affect migration flows only through violence. Two-Stage Least Squares (2SLS, henceforth) method is then performed using these instruments, security expenditure and inequality measures, as well as any combination of them, as they may serve as valid instruments, since they must be uncorrelated to migration flows, except through violence which is included in the equation explaining the dependent variable.

As migration decisions are influenced by observable and unobservable characteristics, it is relevant to control for unobservable issues as well. One way to deal with that is by estimating using fixed effects. Depending on which equation is estimated, you may use different fixed effects. Equation 2.2 is a restricted model where only destination characteristics are important for migration decisions, fixed effects at destination seems to be the most appropriate in this specification. On the other hand, estimating the traditional gravity equation uses destination and origin information. Therefore, it is relevant to control for unobservable characteristics not only at destination but also at origin. Using pair-wise fixed effects might therefore control for all unobservable characteristics at origin and destination which may affect migration flows. Controlling for unobservable characteristics presents a cost: it is not possible to use any time invariant observable

characteristic. For example, this restriction imposes eliminating distance in the traditional gravity equation. As my main variable of interest (violence) varies in time, fixed effects may be interesting to estimate as well.

Another drawback faced by this work, which leads to biased estimated parameters, is the non-existence of migration flows between some regions. The intuition behind this bias is that having information only from non-zero migrations flows, "undesirable" destinations are not considered and their attributes are discarded. Therefore, evaluating migrations decisions using only information of chosen places might lead to biased estimates, since information of non-chosen cities has not been taken into account. Santos Silva and Tenreyro (2006) show an alternative to estimate when zeros are present in a gravity equation. Their methodology is to estimate the econometric specification by a Poisson Pseudo-Maximum Likelihood where the dependent variable remains in its level (not in log) and the explanatory variables are taken in logs. Results from Santos Silva and Tenreyro (2006) show that estimating a gravity equation by Poisson with zeros or not demonstrate no substantial difference. The main reason why truncation has little effect in this case is that observations with zero migration correspond to pairs for which the estimated value of migration is close to zero. The corresponding residuals are hence also close to zero and their elimination from the sample has limited impact.

Summing up, the empirical strategy uses the OLS results as a benchmark. Since violence is an endogenous variable, 2SLS procedure is implemented using government expenditure on security and inequality measures as instruments. As migration may be influenced by unobservable characteristics, fixed effects are performed. The concern of biased estimation due to zero migration flows on a gravity equation is sorted out by using Poison Pseudo-Maximum Likelihood where only non-zero migration flows is taken into

account. [19] All of these methods are used not only when estimating a restricted form, as used in previous studies mentioned earlier, but also by estimating the traditional gravity equation.

2.4. Data Sources

First, sources for such an analysis should be analyzed beforehand. Nunes and Matos (2005) compare two sources of information to analyse migration, where one is based on a sample of formal workers (*RAIS*)[20] and the other based on the Census. [21] One advantage of using Census data is that it has the broadest information since it includes all individuals, from formal to informal workers, all ages, including children and retired people, and so on. Nevertheless, Censuses are taken only every 10 years which limits the number of years. On the other hand, *RAIS* is annual, representing a substantial advantage compared to Censuses, especially because Censuses' data do not contemplate any migration within five years prior to the information collecting year. [22] Nunes and Matos conclusions show that depending on the source of data, there are substantial differences in the results and conclusions. They state that long-distance migrations are more prominent on the Censuses than in *RAIS*. This is an expected outcome since some people move long distances in Brazil to work in the informal sector at their destination. In order to investigate whether violence affects migration, one should take into account not only workers, but also those who work in the informal sector, retired people, and others. Census data seem therefore to be the most appropriate for investigating the issue proposed by this chapter after balancing pros and cons from each source.

[19] Estimation including zeros is not feasible for all information (more than 26 million observations). A random sample of 500 thousand is used however for robustness check.
[20] *RAIS* is *Relacao Anual de Informacoes Sociais* (Social Information Annual Relation), elaborated by the Brazilian Ministry of Labour.
[21] Organized by the Brazilian Statistical Institute (*Instituto Brasileiro de Geografia e Estatística* – [*IBGE*])
[22] Census migration data are related to where people were living five years prior to the Census.

The best data to evaluate this impact are at the individual level, but they are not available. However, regional data can properly answer the research question of this chapter. The focus characteristic (violence) to explain migration affects not only who has been victimized but also relatives and friends. The level of violence in each location influences moreover the whole society by being published regularly in local press, such as radio, newspaper and TV. Even when someone is not related to the victim, he/she might be influenced by what happened to someone else just by being informed about the crime committed. Using regional level data can therefore satisfactorily evaluate whether violence is correlated to migration flows.

Considering which spatial unit better captures these effects, the higher geographical desegregation the better the analysis is, especially since neighbouring effects are investigated. City-level information is the most appropriate regional data available, but shortcomings exist. There were more than one thousand cities emancipated during the 90s. The Brazilian Applied Economical Research Institute (*Instituto de Pesquisa Economica e Aplicada* – [*IPEA*]) has created a classification called comparable minimum areas (*Areas Minimas Comparaveis* – [*AMC*]) by using the city level data from the whole period. This new regional unit comprises 3,559 locations over the period investigated.[23]

Brazilian Population Census provides the main migration flow data for this chapter. From the 1991 and 2000 Censuses, there is information on where people lived five years before the Census year. It is thus possible to build a migration matrix between localities by identifying origin and destination status of urban or rural. Share of urban population uses the same source as migration flows.

For salaries, I use some figures of the economical census from 1985 which includes manufacturing and services. And from 1996, I use the Government Registration of Firms

[23] A detailed description is presented in Appendix 2.I.

(*Cadastro Central de Empresas* [CEMPRE] – from IBGE), which has a limited set of information, such as salaries and number of employees, from all active manufacturing and service firms in Brazil. Although they represent two different data sources, they have exactly the same sample: all manufacturing and service firms in Brazil. Due to data availability, it is possible to distinguish wages into two different sectors: manufacturing and services.[24]

The non-existence of living cost at a much disaggregated level appears as a barrier to create real wages for empirical research. Combes, Duranton and Overman (2005) and Sudekum (2007) pointed out however that housing prices determine living cost when trade freeness is low. This is indeed the case under this study since no trade cost exists (apart from transport cost) within the Brazilian economy. Average housing prices at each location are therefore used as a measure of living cost sourced by IPEA. Even though it looks like a rough proxy, it has been shown theoretically viable, empirically tested by Sudekum (2007) and it is the only data available.

The number of homicides per 100,000 inhabitants in each locality is used as a measure of violence. This information is available from the Mortality Integrated System (*Sistema Integrado de Mortalidade*, acronym SIM) of database of the Brazilian Health Ministry, called Health Database (*Bando de Dados do Sistema Unico de Saude* – DATASUS*)*. Data comprise from 1980 until 2002. Only information about the five years before the census is taken into consideration since it is within this period that people change dwellings. The average of homicide rates over five years before the census data collection year is thus used as a measure of crime rate as well as the level five years before Censuses. Not only violence where it occurs was used in this work, but also its spatial lag by using contiguity matrix or four nearest neighbours, as explained in Section 2.3.

[24] It is also feasible to obtain wages from the agriculture sector, but the focus here is urban-urban migration and this sector does not operate in urban areas.

Controlling for other social amenities, two further variables are utilized as covariates. The number of firms in the service sector, including bars, restaurants, theatres, is used to control for social positive amenities as an entertainment measure. Population density measure is utilized to capture other disamenities, such as congestion and pollution, sourced by IBGE. Additionally, a quadratic in population density is used in order to pick up turning points.

The violence measure may not capture all crimes committed in a city, such as burglaries and thefts; therefore population size may be capturing other types of crime jointly with social network. Summing up, most of the amenities and disamenites will be captured by other measures: natural amenities by salaries; social network and other crimes by population size; social amenity by entertainment measure; and social disamenity, such as pollution and congestion, by density measure. Controlling for all these measures, homicide rates are cleaner to capture only this type of violence.[25]

Percentage of males in each city may also be used to capture any gender issue related to migration, since men are more likely to migrate according to the pools.

Market Potentials are created by using GDP measures provided by IPEA for the period analyzed. Distance was calculated by using the Great Circle Formula from each region centre point.[26] Finally, considering that some export regions might be attractive for migrants, then exporting status is created by assuming that a city which has exported more

[25] Apart from a logical justification, crime rates do not show much correlation within these measures exposed for example, violence and population density were not highly associated since it shows a linear correlation of only 0.24, for the Brazilian case.

[26] Internal distance for Market Potential, according to Redding and Venables (2004), is:

$$dii = (2/3)\sqrt{S_i/\pi}$$

than a certain amount during the years of 2003-04 might have been an exporting region since 1985.[27]

Infrastructure measures are described as follows. For health access, the number of doctors per inhabitant and the percentage of highly qualified nurses catalogued by UNDP are used in the estimation process. The number of years spent on schooling and percentage of illiterate people in each locality is used to control for education infrastructure. Regarding physical infrastructure, the percentage of dwellings with electricity, water supply and sewage is utilized for this purpose.

Distance to the state capital and to the Brazilian capital (Brasilia) are other variables which might be used to capture any movement related to job opportunity in government institutions, since some people move to the capital to start a career as a civil servant. In order to capture the informal sector, the amount of money deposited in the bank might be utilized for this purpose because it is quite common for people working in this sector to have banking or saving accounts. Most of the other variables used in this chapter have also Censuses as a source, such as housing prices, share of urban population, housing infrastructure, *inter alia*. A full description of the variables used in this chapter is listed in Appendix 2.I.

2.5. Econometric Results

Only places, as stated previously, where migration between two localities was greater than zero are taken into account for econometric results in this section. Even though this restriction on the data reduces the sample substantially, there are still more than 100

[27] The threshold used to create this dummy for exporting status was US$ 500 billion. Other measures were used, such as US$ 1 trillion and US$ 1 billion, but results remained practically similar.

thousand valid observations for the two years investigated.[28] Poisson pseudo-maximum likelihood does not present a substantial difference when estimating with or without zero, as discussed earlier.[29] Most of the explanatory variables are normalized by national figures.[30] An unemployment rate of a region *j*, for example, is divided by the Brazilian unemployment rate. This avoids some national time trend (such as recession and/or inflation) which may be associated with changes, but it does not remove time changes between localities.[31] However, time dummy is also included to pick up any other time trend not presented in the observable variables. Before presenting any outcome, it is important to mention that standard errors are robustly estimated throughout this whole section.

Table 2.3 shows results from the restricted model presented in Equation 2 estimated by OLS (1st column), 2SLS (2nd and 3rd columns),[32] Destination Fixed Effects (4th column) and Poisson Pseudo-Maximum Likelihood (last column).

[28] Some regions have no doctor; therefore it is not possible to take log. This observation is therefore discarded. Levels, instead of log of the variable, are also used for estimation to overcome this issue and similar results emerge. See Appendix 2.II.

[29] Results from a random sample of 500 thousand migration flows from all observations, including zeros and non-zeros, are presented for robustness check in Appendix 2.II.

[30] Exceptions are those time invariant, such as distances and dummy variables.

[31] Results using pair-wise violence differentials were also estimated and outcomes just corroborate results presented in this section. However, they cannot identify whether violence is a pulling factor at origin and/or a non-pushing variable at destination. Therefore, having variables identifying levels of violence at origin and destination help to understand how crime rates might affect people's choice in a migration context.

[32] The second column shows results using only the percentage of state expenditure on security while the third column presents outcomes using the previous instrument added to inequality measures (Gini and Theil) and other death rates (suicide rate and death victims in car accidents).

Table 2.3: Restricted Model Results

Panel A: Regression Results					
Dependent Variable	(i)	(ii)	(iii)	(iv)	(iv)
Share of Migrants	OLS	2SLS	2SLS	FE Dest.	Poisson
Violence	0.04	-0.712	-0.461	0.007	-0.085
	(0.016)**	(0.196)***	(0.158)***	(0.023)	(0.021)***
Pop. Density	0.079	-0.257	-0.145	-0.525	-0.136
	(0.125)	(0.194)	(0.165)	(0.197)***	(0.154)
Pop. Density Sqr.	0.004	-0.024	-0.015	-0.052	0.014
	(0.006)	(0.012)**	(0.01)	(0.019)***	(0.006)**
Entertainment	0.023	0.154	0.111	0.086	0.176
	(0.044)	(0.087)*	(0.071)	(0.057)	(0.058)***
Share of Urban Pop	-6.567	-11.687	-9.98	1.073	-2.415
	(0.843)***	(2.312)***	(1.868)***	(2.153)	(1.170)**
Distance	-0.739	-0.708	-0.718	-0.772	-0.623
	(0.030)***	(0.032)***	(0.032)***	(0.033)***	(0.058)***
Manuf. Real Wage	-0.095	0.001	-0.031	0.068	-0.149
	(0.027)***	(0.049)	(0.041)	(0.035)*	(0.040)***
Service Real Wage	0.028	0.349	0.242	-0.007	0.064
	(0.032)	(0.108)***	(0.085)***	(0.033)	(0.043)
Unemployment Rate	0.031	0.226	0.161	-0.053	0.131
	(0.046)	(0.098)**	(0.080)**	(0.066)	(0.065)**
Market Potential	-0.121	0.236	0.117	0.974	-0.342
	(0.032)***	(0.110)**	(0.088)	(0.282)***	(0.054)***
GDP Growth	0.004	-0.022	-0.014	0.006	0
	(0.01)	(0.021)	(0.017)	(0.009)	(0.011)
Distance to State Capital	-0.036	-0.11	-0.085	-	-0.047
	(0.012)***	(0.030)***	(0.024)***	-	(0.014)***
Distance to Brasilia	-0.077	-0.022	-0.04	-	-0.047
	(0.022)***	-0.029	(0.018)**	-	(0.020)**
Electricity	-0.894	-0.419	-0.577	0.239	-0.794
	(0.211)***	(0.324)	(0.275)**	(0.324)	(0.249)***
Water Supply	0.032	-0.278	-0.174	-0.302	-0.06

	(0.078)	(0.131)**	(0.108)	(0.098)***	(0.089)
Sewage	-0.032	-0.033	-0.032	0.005	0.001
	(0.009)***	(0.015)**	(0.012)***	(0.009)	(0.01)
Educational Level	0.124	1.136	0.799	0.408	0.232
	(0.157)	(0.386)***	(0.306)***	(0.335)	(0.205)
Illiteracy	0.12	0.737	0.532	-0.078	0.152
	(0.060)**	(0.196)***	(0.159)***	(0.156)	(0.072)**
Informal Sector	0.036	-0.013	0.003	-0.006	0.006
	(0.015)**	(0.032)	(0.025)	(0.022)	(0.019)
Share of Male	7.008	12.31	10.543	-0.899	2.911
	(0.806)***	(2.308)***	(1.862)***	(2.131)	(1.159)**
Doctors	-0.043	-0.059	-0.053	-0.005	0.004
	(0.016)***	(0.032)*	(0.025)**	(0.02)	(0.02)
Share of Manuf Sector	0.062	-0.042	-0.008	-0.053	0.105
	(0.017)***	(0.039)	(0.032)	(0.029)*	(0.024)***
Share of Service Sector	0.003	0.002	0.002	0.034	-0.152
	(0.041)	(0.075)	(0.058)	(0.05)	(0.046)***
Observations	150,992	150,992	150,992	150,992	150,992
R-squared	0.32	0.25	0.29	0.34	0.10
Panel B: First Stage Results					
Instruments used	-	Security Expenditure	Sec. Expenditure & Gini Index	-	-
First Stage R-squared	-	0.03	0.04	-	-
First Stage F-stat.	-	4,401	2,835	-	-
Over-identification test	-	-	12.6	-	-

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Before focusing on the interested variables (amenities), an overview of other variables is essential. In general, results show the majority of the variables are relevant to explain migration flows, since most of the estimated parameters are significant,[33] but not all of them have the expected sign. While distance presents the negative and significant sign in

[33] Similar results are found for example in Poncet (2006) for China case.

all procedures, market potential does not show robust results. This result about market potential might be partly explained by the fact that real wages are included in the econometric specification. Wages in services seem to attract migrants because it has a positive and significant value in 2SLS, but non significant in other methods. Manufacturing real wages however mainly present the wrong sign, apart from fixed effects approach. Similar outcomes are found nevertheless in Crozet (2004) in which services tend to explain more internal migration than manufacturing sector.

Some other variables show interesting results. Distance to state and federal capitals also present very robust results showing the closer a city is to any capital, the more people migrate to these localities.[34] This may represent some migration related to labour opportunities in state or federal governments, as mentioned earlier, where people move to the capital after being enrolled as a civil servant. The economic performance of a region, measured by the increase in GDP, seems not to matter for migration flows, because no association between them is found. Apart from fixed effects approach, urban size of the destination presents a negative and significant sign which corroborates findings by Andrade et al. (2000) on the increased importance of medium-sized cities in an urban population.[35] This outcome may be interpreted as the other non controlled negative amenities (such as pollution) are greater than any other positive one, social network for example. People may have tended to avoid large cities for the period analyzed.

Entertainment amenity and population density (representing a disamenity) present expected signs, but results are not robust. While entertainment amenity appear positive and significant only in Poisson and 2SLS (using security expenditure as instrument), population density is negative and significant solely in FE at destination. Those results suggest that

[34] As distances to state and federal capital are time invariant, they drop when fixed effects at destination is included.
[35] Urban size is endogenous to migration flows. However, migration flows are small compared to total population.

people seem to be avoiding more dense areas (where congestion is more evident) and willing to live in areas with more entertainment options. Although those results are not robust, they might provide us some weak evidence that social amenities might have a correlation to where people decide to live.

Our focus variable, violence, has a positive and significant sign in the OLS results, but the parameter estimated might be biased. It is hence important to instrument it to see whether this positive association changes. After using instruments in the 2SLS, it is possible to notice that violence becomes negative and significant, despite the number of instruments are used. Nevertheless, some more diagnosis should be done in order to evaluate the adequacy of those instruments. Some conditions should be met, which are: instruments should be sufficiently correlated with the endogenous variable, which is violence; instruments should meet exclusion restriction.

First Stage results of IV procedure present some evidence whether instruments are statistically correlated to the endogenous variable. Security measure can explain only a small part of the variability of the endogenous variable, since R-square of the first stage is only 0.03. This is partly explained by the fact that instrument is at state level, while violence is at city level. When other instruments of city level are included, available at column (iii), R-squared increases slightly. Nevertheless, instruments are statistically relevant to explain violence since p-value of the t-statistic of the parameters estimated is significant at any level. F-statistic of first stage backs up those findings as shown in Table 2.3. Staiger and Stock (1997) formalized the definition of weak instruments, but researchers have found that any F-statistic in the first stage exceeding 10 would lead to the conclusion that instruments are sufficiently strong. Therefore, instruments used in this work would be considered strong following Staiger and Stock (1997) definition, since F-statistic are far beyond the threshold mentioned.

However, this is not sufficient to establish that instruments are valid, since orthogonality condition should be also met. Moreover, if there is more than one instrument, however, two-stage least squares provides asymptotically efficient estimates of the parameter on the endogenous independent variable compared to a just identified IV estimation. As it is known, when the equation is just identified, such as in the second column, it is not possible to perform over-identification tests. Then, security expenditure and Gini index at municipality level create a set of instruments to perform over-identification tests, as shown in the third column of Table 2.3. Since the number of instruments exceeds the number of endogenous regressors, 2SLS is over-identified, allowing for a test of the exogeneity of this instruments' set. The tests of over-identifying restrictions yield more disappointing results. Over-identification tests show instruments fail to meet all conditions necessary for instruments' validation, since p-value of a chi-square at 12.6 with 1 degrees of freedom is nearly zero. Different sets of instruments have presented similar result, such as using other inequality measure (Theil index) and also other spatial scale for Gini index (state level). Summing up, all sets of instruments used fail on over-identification tests. One possible explanation is that instruments might be affecting the dependent channels apart from the violence measure used in this estimation process. Another feasible justification might be reverse causality. In other words, perhaps violence may have been caused by migration flows. Additionally, those outcomes also raise question on whether security expenditure also meet the orthogonality condition, although it is not possible to perform over-identification test with this instrument solely. As a matter of simplicity, outcomes using only security expenditure as instrument are shown in further tables, since this variable presents some economic intuition to act as an instrument.

As IV results are not reliable and other issues should be addressed, two further methods are implemented. Fixed effects present different results from OLS and IV.

Controlling for unobservable characteristics at destination, violence becomes non-significant to explain migration flows. In the last column, Poisson Pseudo-Maximum Likelihood results show that violence is negative and significant, which might suggest that people might moving towards cities with lower crime rates.

Considering this restricted model, outcomes suggest violence may be relevant in explaining to where people migrate: cities with lower crime rates tend to attract more migrants than their counterparts. The migration literature suggests that fear of crime leads people to move towards medium-sized cities, even though the level of crime could be similar between origin and destination. Brazilian case presents however that the level of violence may represent these fears, since results suggests that it is negatively correlated to migration flows. It is thus not only a perception of violence, but actually their values which are correlated to people's decision to change dwelling across cities.

One may argue that migration flows might have distinct patterns depending on how far the two localities are apart. Housing prices, for example, might be much more relevant for changing dwellings between cities which are less than a certain limit than for those beyond this threshold. On the other hand, long-distance migration flows might be more related to differences in nominal wages than short ones, because people can change job from one city to another but not dwelling. As it regards amenities, especially violence, there might be some differences as well because migration flows might be more strongly correlated at short distances than long ones, since people may change only where to live and not where they work. Following Table 2.3 format, Table 2.4 presents results[36] splitting by migration below and beyond 100 km[37] isolating the effect of housing prices and nominal

[36]Columns (i) and (v) present OLS results. Columns (ii) and (vi) present 2SLS using only Security as instrument. Columns (iii) and (vii) show FE at destination. Finally, columns (iv) and (viii) show Poison pseudo-maximum likelihood outcomes.

[37]One hundred kilometres is chosen as a threshold for two reasons. First, 95% of the distances between contiguous cities are below 75 km. Second, 60% of the migration flows occurs underneath 250km, where half

wages. For parsimony, results only for the most relevant variables in this study are presented in Table 2.4.

Table 2.4: Restricted Model Results by Short and Long Migration Flows

Panel A: Regression Results								
Dependent Variable	Under 100km				Beyond 100km			
Share of Migrants	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Violence	-0.341	-0.602	0.02	-0.257	0.07	-0.603	0.00	-0.057
	(0.033)***	(0.115)***	(0.035)	(0.033)***	(0.019)***	(0.170)***	(0.026)	(0.030)*
Pop. Density	-0.698	-0.654	-0.728	-1.344	0.081	-0.039	-0.288	-0.293
	(0.328)**	(0.323)**	(0.371)**	(0.298)***	(0.137)	(0.201)	(0.203)	(0.188)
Pop. Density Sqr.	-0.021	-0.02	0.014	-0.059	-0.009	-0.021	-0.052	-0.009
	(0.015)	(0.016)	(0.035)	(0.014)***	(0.006)	(0.010)**	(0.021)**	(0.005)*
Entertainment	0.336	0.336	0.23	0.345	-0.004	0.087	0.078	0.056
	(0.080)***	(0.081)***	(0.083)***	(0.073)***	(0.051)	(0.091)	(0.066)	(0.073)
Share of Urban Pop	3.935	1.644	-0.239	4.181	-8.521	-14.89	0.585	-5.351
	(1.503)***	(1.913)	(2.181)	(1.408)***	(0.891)***	(2.416)***	(2.238)	(1.480)***
Distance	-1.199	-1.202	-1.372	-0.91	-0.612	-0.563	-0.662	-0.462
	(0.026)***	(0.026)***	(0.025)***	(0.030)***	(0.038)***	(0.043)***	(0.043)***	(0.093)***
Manuf. Nominal Wage	-0.215	-0.178	-0.069	-0.33	-0.082	0.061	0.097	-0.148
	(0.052)***	(0.053)***	(0.052)	(0.057)***	(0.033)**	(0.06)	(0.039)**	(0.052)***
Service Nominal Wage	-0.045	0.03	-0.049	-0.031	0.09	0.328	0.012	0.085
	(0.057)	(0.067)	(0.06)	(0.064)	(0.041)**	(0.093)***	(0.041)	(0.059)
Housing Prices	-1.178	-1.318	-0.189	-1.076	0.264	0.103	0.163	-0.143
	(0.188)***	(0.189)***	(0.131)	(0.186)***	(0.093)***	(0.165)	(0.097)*	(0.14)
Observations	26,067	26,067	26,067	26,067	124,925	124,925	124,925	124,925
R-squared	0.37	0.36	0.55	0.10	0.27	0.21	0.31	0.07
Panel B: First Stage Results								
R-squared	-	0.05	-	-	-	0.03	-	-
F-stat.	-	1,475	-	-	-	4,182	-	-

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

of them below 50km. Thus, this limits of 100km represents not only a reasonable amount of the migration flows but also migrants who may have changed residence but not job.

Regarding the main difference in the explanatory variables from the first econometric results, nominal wages in the service sector seems to matter more for long distance migration flows than for short distance. Manufacturing nominal wages appear again not presenting robust results. House prices below 100km present negative expected sign, apart from destination fixed effects, while beyond this threshold only Poisson shows the negative expected sign. This result corroborates the finding of Jannuzzi and Jannuzzi (2002) in which housing price is the main factor to explain migration in the Sao Paulo metropolitan region. Overall, wages and housing prices present expected outcomes, which the latter seems to be more relevant for short distance, while the former for long distance. Distance still plays an important role regardless of whether migration flows are short or long distance.

One important difference is regarding urban population size. Although migration flows are correlated to less populated places beyond 100 km, the contrary occurs underneath this threshold. This means that people in short distances tend to move towards bigger cities, while in long distance they tend to migrate towards smaller cities. The strength of social network is greater than the negative amenities when cities are mainly neighbours while the opposite happens in long distance migration flows. One explanation for this outcome may be that people who are in surrounding cities may be also affected by these negative amenities at their origin cities; then moving towards a bigger city seems not bother them.

As it regards violence, outcomes present pretty similar results which mean that people are moving towards safer places.[38] Comparing the strength of this association, it is possible to infer that a greater correlation is found at a short distance than at long by looking at Poisson results. Considering these results, crime rates at short distance is nearly

[38] Results from IV estimation under 100 km should be analyzed with care, since security measure is at a state level. Results on this issue have therefore become positive, perhaps because this instrument does vary substantially under this threshold.

twice correlated than those at long distances.[39] Violence affects more migration flows under 100 km than beyond this threshold. This is an expected outcome since some migrants are able to move only their residence (from a violent city towards a safer one) rather than both, job and dwelling, in long distances.

Results obtained so far do not consider origin characteristics. Some part of the literature argues that what really matters is destination attributes because people do not value the characteristics of origin since the decision to move is already made. The destination attributes are therefore much more relevant to evaluate than origin. However, the variable investigated here, violence, is a disamenity where origin might be more relevant than destination. In other to check whether more violent cities tend to repel people, a traditional gravity equation is estimated for this purpose. Some slight differences emerge for this exercise as stated previously. First, the dependent variable cannot be the share of migrants but rather the total migrants. In the end, the dependent variable turned out to be a percentage of migration *i* to *j* in the total migration flows of year *t*, as explained in Section 2.3. Second, all variables which are used at destination should also be utilized for origin. In order words, share of population, for example, from origin and destination is included in the estimation process. Third, as origin variables are also included, and then fixed effects should consider not only unobservable destination characteristics but also origin ones. In order to consider both, pair-wise unobservable attributes are controlled in the fixed effects approach.[40] Table 2.5 shows a summary of the results of a traditional gravity equation.[41]

[39] Considering the 95% confidence interval, the upward limit of long distance is around -0.11 and the downward limit for short distance is -0.19.
[40] FE has created more than 100 thousands pair-wise, leaving around 10 thousand observations which are used to estimate parameters needed.
[41] The full table is presented at Appendix 2.II.

Table 2.5: Traditional Gravity Equation Results

Panel A: Regression Results					
Dependent Variable	(i)	(ii)	(iii)	(iv)	(iv)
Share of Migrants	OLS	OLS	2SLS	FE Pair-wise	Poisson
Distance	-0.317	-0.417	-0.413	-	-0.889
	(0.002)***	(0.003)***	(0.004)***	-	(0.014)***
Share of Pop.	0.296	-2.622	-2.911	-1.388	-11.495
	(0.001)***	(0.185)***	(0.212)***	(1.307)	(0.625)***
Share of Pop. Origin	0.213	1.273	1.308	4.395	4.153
	(0.002)***	(0.183)***	(0.197)***	(1.253)***	(0.717)***
Violence		0.011	-0.085	-0.037	0.004
		(0.004)***	(0.028)***	(0.020)*	(0.011)
Violence Origin		0.062	0.053	0.052	0.097
		(0.004)***	(0.020)***	(0.019)***	(0.012)***
Pop. Density		-0.025	-0.076	-0.642	0.026
		(0.025)	(0.028)***	(0.160)***	(0.083)
Pop. Density Sqr.		-0.011	-0.012	-0.024	-0.007
		(0.001)***	(0.001)***	(0.017)	(0.004)*
Entertainment		0.11	0.111	0.024	0.324
		(0.010)***	(0.010)***	(0.051)	(0.039)***
Pop. Density Origin		0.223	0.219	0.35	0.306
		(0.028)***	(0.028)***	(0.187)*	(0.086)***
Pop. Density Sqr. Origin		0.016	0.016	0.03	0.039
		(0.002)***	(0.002)***	(0.019)	(0.006)***
Entertainment Origin		-0.039	-0.042	0.08	0
		(0.010)***	(0.010)***	(0.05)	(0.03)
Observations	275,308	113,307	113,307	113,307	113,307
R-squared	0.26	0.28	0.28	0.02	0.06
Panel B: First Stage Results					
Instruments	-	-	Security Exp.		
R-squared (Dest / Origin)	-	-	0.04 / 0.06	-	-
F-stat. (Dest / Origin)	-	-	2,152 / 3,567	-	-

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

In order to show that control variables do not change their results substantially, distance remains significant and negative. Estimating only with distance and urban population size, there is no clear evidence whether urban population size at origin and destination have different signs. Only with OLS urban population does have the expected significant sign at destination, but not at origin. After controlling for all other characteristics, urban population turns out negative at destination and positive at origin in all methods analyzed. Negative amenities appear to overcome social network even after controlling for origin attributes since it has a significant and negative sign at destination. Looking at both results (origin and destination), it is feasible to infer that these outcomes corroborate once more the stylized fact that people are moving from metropolitan cities towards medium-sized cities.

Entertainment amenity presents a significant expected sign at destination and negative at origin only when OLS, 2SLS and Poisson are implemented. This outcome represents the traditional behaviour of gravity equation where origin attributes are centrifugal forces towards destination with a centripetal one. Population density, used as a proxy for congestion, present some results (2SLS and FE) at destination showing that people are avoiding densely populated regions. On the other hand, this disamenity has a positive sign at origin in all methods implemented, which means that this attribute may be more relevant as a pushing factor at origin than a pulling factor at destination.

Violence now seems to act much more as a pushing factor at origin than as a pulling factor at destination. Crime rates at destination do not present significant negative results only by using Poisson, which might be a reasonable indication that people are moving towards cities with lower violence records. Crime rates at origin show however significant positive outcomes in all methods. These findings suggest the level of violence at origin might be more important in explaining migration flows than its level at destination.

Combining with the urban population result, it is feasible to conclude that metropolitan areas where crime rates occur more often are definitely repelling people towards smaller and safer places, which are medium-sized cities.

As discussed previously, neighbouring effects of violence might be relevant in explaining migration flows because violence affects people's well being not only locally but in surrounding areas as well. The main explanation is due to social interaction of violence. Table 2.6 shows a summary of the results with spatial lag variable using the contiguity matrix.[42] The first four columns present results using the restricted gravity equation while the last four show outcomes using the traditional gravity equation.

Apart from FE in the restricted equation, urban population continues showing for example the significant negative sign at destination. Positive and negative amenities do not change their previous interpretation after the inclusion of this new variable disregard the model adopted.

As it regards violence, it is important to mention that the inclusion of spatial lag remove the negative significant sign of local crime rates in all methods considering the restricted equation. The spatial lag do not show robust result for security as a pulling factor at destination, since only in the Poisson using the restricted equation. Outcomes change however at the traditional gravity equation. Once more, crime rates at origin appear to act as a pushing factor for migrants much more robustly than as a pulling one at destination, not only locally but also in neighbouring areas. This corroborates the fact that violent places or those surrounded by areas with high levels of criminality are repelling people to other places, which might have lower offences registered at police stations.

[42] For simplicity, only results of amenities and distance are shown in Table 2.6. Additionally, IV column means instrumenting only by security measure. All regression uses real wages. A table with all results is available at Appendix 2.II. The instruments for the spatial lag of violence are the spatial lag of security expenditure.

Table 2.6: Neighbouring Effects Using Contiguity Matrix

Panel A: Regression Results								
Gravity Equation Type	Restricted Gravity Equation				Traditional Gravity Equation			
Dep. Var.: Share of Migrants	OLS	IV	FE	Poisson	OLS	IV	FE	Poisson
Violence	0.06	0.66	-0.01	-0.04	0.01	-0.39	-0.04	-0.01
	(0.023)**	(0.82)	(0.03)	(0.03)	(0.01)	(0.228)*	(0.024)*	(0.02)
Spatial Violence Contiguity	-0.03	-0.91	0.01	-0.07	0.00	0.24	-0.01	-0.01
	(0.02)	(0.65)	(0.03)	(0.030)**	(0.01)	(0.17)	(0.03)	(0.02)
Violence Origin					0.04	-0.97	0.02	0.06
					(0.006)***	(0.572)*	-0.02	(0.018)***
Spatial Vio Origin Contiguity					0.03	0.97	0.07	0.06
					(0.006)***	(0.558)*	(0.024)***	(0.019)***
Share of Pop	-6.57	-11.57	0.75	-2.42	-2.59	-1.92	-2.37	-11.40
	(0.891)***	(2.864)***	(2.227)	(1.216)**	(0.196)***	(0.551)***	(1.387)*	(0.645)***
Share of Pop Origin					1.40	4.42	4.27	3.87
					(0.196)***	(1.837)**	(1.320)***	(0.695)***
Pop. Density	0.11	0.29	-0.42	-0.11	-0.04	-0.17	-0.70	-0.04
	(0.12)	(0.34)	(0.201)**	(0.16)	(0.03)	(0.064)***	(0.170)***	(0.08)
Pop. Density Sqr.	0.01	0.01	-0.04	0.02	-0.01	-0.02	-0.04	-0.01
	(0.00)	(0.02)	(0.020)**	(0.005)***	(0.001)***	(0.002)***	(0.018)**	(0.004)**
Entertainment	0.00	-0.03	0.07	0.14	0.11	0.10	-0.01	0.32
	(0.04)	(0.13)	(0.06)	(0.055)**	(0.011)***	(0.016)***	(0.05)	(0.039)***
Pop. Density Origin					0.22	0.38	0.38	0.31
					(0.029)***	(0.103)***	(0.198)*	(0.086)***
Pop. Density Sqr. Origin					0.02	0.05	0.04	0.04
					(0.002)***	(0.019)**	(0.020)*	(0.006)***
Entertainment Origin					-0.04	0.00	0.06	0.00
					(0.011)***	(0.03)	(0.05)	(0.03)
Observations	141,727	141,727	141,727	141,727	101,270	101,270	101,270	101,270
R-squared	0.32	0.26	0.34	0.10	0.28	0.14	0.03	0.05
Panel B: First Stage Results								
R-Square (Destination)	-	0.03	-	-	-	0.04	-	-
F-Stat. (Destination)	-	2,488	-	-	-	1,132	-	-
R-Square (Origin)	-	-	-	-	-	0.06	-	-
F-Stat. (Origin)	-	-	-	-	-	1,753	-	-
R-Square (Spatial Dest.)	-	0.06	-	-	-	0.07	-	-
F-Stat. (Spatial Dest.)	-	4,357	-	-	-	1,858	-	-
R-Square (Spatial Origin)	-	-	-	-	-	0.07	-	-
F-Stat. (Spatial Origin)	-	-	-	-	-	1,934	-	-

(a) Robust standard errors in parentheses; (b) * significant at 10%; ** significant at 5%; *** significant at 1%

As discussed in Section 2.3, one possible criticism is that not only contiguous municipalities should be considered but also their closest neighbours, which may not be contiguous. In order to tackle this issue, Table 2.7 shows results using the four nearest neighbours' matrix to construct the spatial lag violence.[43]

Urban population continues to have the expected outcome. Population density and entertainment amenity show less results from before although the same interpretations, which is they tend to expel people from places with more congestion and less entertainment options towards others with the opposite values.

Regarding violence itself, the neighbouring effects also change violence findings where the spatial lag at destination seems to be more negatively correlated to migration flows than at the local level. In the traditional gravity equation, crime rates at origin, locally or spatially, appear to continue to be positively associated to outward migrants. In order words, these findings suggest that criminality figures from where people are moving from seem to be positively correlated to migration flows not only locally but also spatially. On the other hand, places surrounded by safer municipalities appear to be attracting more inhabitants than those with lower criminal records solely.

Considering both matrices to construct the spatial lag, overall findings suggest that neighbouring effects seem to be associated to migration flows and they do not eliminate the effects of local violence, especially at origin.

[43] A table containing all parameters estimated is presented in the Appendix 2.II.

Table 2.7: Neighbouring Effects Using Four nearest Neighbours Matrix

Panel A: Regression Results								
Gravity Equation Type	Restricted Gravity Equation				Traditional Gravity Equation			
Dep. Var.: Migrants Share	OLS	IV	FE	Poisson	OLS	IV	FE	Poisson
Violence	0.09	6.50	0.00	0.04	0.01	1.56	-0.04	0.03
	(0.024)***	(17.69)	(0.03)	(0.03)	(0.005)*	(2.98)	(0.022)**	(0.02)
Spatial Violence Nearest	-0.06	-4.37	0.01	-0.15	0.00	-1.02	0.00	-0.04
	(0.022)***	(10.54)	(0.03)	(0.029)***	(0.01)	(1.85)	(0.02)	(0.020)**
Violence Origin					0.04	1.67	0.03	0.06
					(0.005)***	(3.39)	(0.02)	(0.020)***
Spatial Vio Origin Nearest					0.04	-1.66	0.04	0.07
					(0.006)***	(3.42)	(0.022)*	(0.021)***
Share of Pop	-6.66	2.50	1.22	-2.82	-2.59	-2.68	-1.28	-11.66
	(0.839)***	(40.423)	(2.147)	(1.119)**	(0.185)***	(0.513)***	(1.32)	(0.633)***
Share of Pop Origin					1.41	-3.78	4.58	4.46
					(0.184)***	(10.11)	(1.262)***	(0.729)***
Pop. Density	0.08	2.13	-0.54	-0.10	-0.03	0.40	-0.61	0.04
	(0.13)	(5.85)	(0.198)***	(0.15)	(0.03)	(0.86)	(0.162)***	(0.08)
Pop. Density Sqr.	0.00	0.16	-0.05	0.02	-0.01	0.00	-0.02	-0.01
	(0.01)	(0.46)	(0.019)***	(0.006)***	(0.001)***	(0.03)	(0.02)	(0.004)*
Entertainment	0.02	-1.00	0.09	0.15	0.11	0.07	0.03	0.32
	(0.04)	(2.91)	(0.06)	(0.054)***	(0.010)***	(0.08)	(0.05)	(0.039)***
Pop. Density Origin					0.24	-0.26	0.36	0.32
					(0.028)***	(0.99)	(0.189)*	(0.085)***
Pop. Density Sqr. Origin					0.02	-0.04	0.03	0.04
					(0.002)***	(0.13)	(0.02)	(0.006)***
Entertainment Origin					-0.04	-0.07	0.08	0.00
					(0.010)***	(0.09)	(0.05)	(0.03)
Observations	150,483	150,483	150,483	150,483	112,661	112,661	112,661	112,661
R-squared	0.32	0.27	0.34	0.10	0.28	0.15	0.03	0.06
Panel B: First Stage Results								
R-Square (Destination)	-	0.03	-	-	-	0.04	-	-
F-Stat. (Destination)	-	2,279	-	-	-	1,153	-	-
R-Square (Origin)	-	-	-	-	-	0.06	-	-
F-Stat. (Origin)	-	-	-	-	-	1,987	-	-
R-Square (Spatial Dest.)	-	0.06	-	-	-	0.08	-	-
F-Stat. (Spatial Dest.)	-	5,203	-	-	-	2,477	-	-
R-Square (Spatial Origin)	-	-	-	-	-	0.08	-	-
F-Stat. (Spatial Origin)	-	-	-	-	-	2,501	-	-

(a) Robust standard errors in parentheses; (b) * significant at 10%; ** significant at 5%; *** significant at 1%

2.6. Conclusion

This chapter investigates whether amenities/disamenities, especially violence, are correlated with urban-urban migration flows between Brazilian cities. Not only violence was used to explain it, but also other amenities and disamenities, such as population density and entertainment varieties as well as many social and economic variables.

Amenities investigated in this chapter present interesting results. Entertainment amenities seem to be an important pulling factor for migration flows which means that people do migrate towards cities with larger options of restaurants, shops, and others. Disamenities, such as congestion, appear to be the pushing force representing people's desire to move to less densely populated cities. Urban population outcomes corroborate Andrade et al. (2000) findings in which medium-sized cities are increasing their participation in the Brazilian urban population since larger places are repelling inhabitants towards smaller cities. Therefore, the negative part of larger areas, such as pollution and other crimes, appears to overcome most of the positive part, social network, for example.

Looking at violence, the results suggest it is negatively correlated to Brazilian migration flows. Places offering lower levels of crime rates tend to have more people moving towards them. Neighbouring effects are also tested and outcomes show cities with high violence in their surrounding territories do not attract migrants. Additionally, cities surrounded by violent places also repel people towards other places. It is therefore safe to conclude that it is not possible to exclude criminality to explain the movement of people throughout all of the Brazil.

Chapter 3

Trade Shocks in Brazil: An Investigation of Effects on Regional Manufacturing Wages

In the previous chapter, individuals' decision over where to live is examined. People move only to where employment opportunities arise, which are offered by firms. Firms' location decision fills this gap. This chapter explores some issues related to firms' location.

In order to maximize profits, one decision which firms face is where to locate their production plant. The New Economic Geography (NEG hereafter) is a branch of the economic literature which investigates this dilemma. According to this literature, a negative correlation exists between two marginal costs: transport cost; employee's wages. In other words, when transport cost rises, salaries reduce. Proximity to markets explains this negative association, since firms face lower transport cost to ship their goods to consumers when their production plant is close to the market. Firms located far from markets need therefore to offer lower wages to their employees in order to compensate a higher transport cost. This trade-off faced by any firm promotes regional wage disparities.[44]

This pattern may change however after opening to trade, since the importance of external and internal market changes. This chapter investigates these features by testing how demand linkages are important to explain regional wage disparities. Additionally, it also investigates whether a trade shock is able to influence these inequalities by also trying to measure which reduction in trade cost (export or import) appears to be stronger.

[44] Actually, the role of transport cost in determining location is older than NEG. Fujita, Krugman and Venables (1999) list some earlier work. Von Thünen (1826) model is an example. It explains the negative relation between land rent and distance to a city. Other more recent examples are from the Central Place Theory, such as Christaller (1933) and Lösch (1940), in which they examine the relation between economies of scale and transport costs.

Agglomeration and dispersion forces form the basis for understanding location decisions, especially after a reduction in trade costs. The main example of agglomeration force is called as "home market effect" by Krugman (1980). Regarding dispersion forces, the main one evidenced by the literature is the increase competition in the market. When trade costs reduce, foreign market tends to increase importance for exporting firms due to more demand for their final goods. Firms importing inputs from abroad may find regions close to external market more attractive because they are able to buy cheaper inputs. For those two types of firms, the importance of the foreign market increases relatively to domestic market after a reduction of trade costs. Wages in those regions therefore tend to close the gap with the industrial center.[45] Wages inequalities between regions may eventually diminish. Nevertheless, some other firms might not perceive this change in international trade costs the same way. Firms facing higher competition with import goods perceive this decrease in trade cost as a threat to their survival, since external products become available at a lower cost. These firms therefore might tend to locate even farther from the external market.[46] In either case, regional wage disparities may therefore change after manufacturing geographical production relocation due to a trade shock.

The reason why regional wages might be influenced by a reduction in trade costs can be explained by the fierceness in competition in the goods market and in the labor market. On one side, cheaper imported goods force existing firms within a market to reduce their price; those not competitive enough either close down or locate elsewhere to survive. Therefore, regions close to external market are more affected by this shock, where firms lower employee's wages to compete with cheaper imported goods. On the other hand, greater access to the external market stimulates firms to hire extra labor force in order to

[45] Assuming that the industry centre is not close to the external market, as highlighted by Henderson, J. V. (1996).
[46] Industry centre is an option, but not quite, since these firms could not survive there before trade cost reduction.

expand production. Demand for labor rises and eventually so do wages. Regions located close to the external market face higher competition on their demand for labor force and less efficient firms have two options: either exit the market or locate elsewhere. Therefore, non competitive firms have only two options after facing more competition: either exits the market or change location. Both attitudes influence regional wages; however the strength of each dispersion force has not been investigated in the literature. In a trade agreement, these two different channels occur simultaneously. Reduction of trade cost raises competition with imports; likewise, exporting firms may expand their production, increasing demand for labor. The result therefore might even be neutral since jobless employees from firms facing higher competition with imported goods can be hired by exporting firms expanding their production.

There are two venues explored by the literature to address how demand linkages can affect regional wage disparities. One part of it investigates those linkages by using transport cost as the main explanatory variable for regional disparities in wages, such as Hanson (1996; 1997) and Brulhart and Koenig (2006). Another more recent and numerous part of the literature, represented by Brakman, Garretsen and Schramm (2004); Mion (2004); Redding and Venables (2004); Hanson (2005); Head and Mayer (2006); Fally, Paillacar and Terra (2008), tests how important market potential is to explain those spatial inequalities.[47] Hanson (1996; 1997) investigates the Mexican case by not only exploring how transport cost over Mexican states may explain regional wage disparities, but also by probing whether the North America Free Trade Agreement (NAFTA) had any impact on these regional inequalities. Apart from Hanson (1996; 1997), the remaining papers mentioned earlier do not address whether trade shocks may impact regional wages, including Fally et al (2008) over Brazil. This chapter contributes to this literature by

[47] These two ways (transport cost and market potential) are not exclusive, but rather the same. In other words, the lower the distance is to markets (which is the higher the market potential), the higher are wages.

showing not only how both, transport cost and market potential, can explain regional wage differences in Brazil, but also how trade shocks affect these disparities.

Brazil provides a good example to investigate these phenomena. First, this country had a closed economy until the late 80s. During the 90s, it experienced two trade shocks which reduced trade costs: a unilateral liberalization process which diminished the weighted average nominal import tariff from 37.7% in 1988 to 10.2% in 1994; and a dramatic exchange rate devaluation of 47% in 1999.[48] Additionally to these shocks, there was a stabilization plan for hyper-inflation in 1994 which has appreciated the Brazilian currency (named "Real") against others. This was an extra macroeconomic shock in the local economy, which has also impacted the Brazilian economy by providing cheaper imported goods.

These shocks have substantially affected Brazilian trade flows. After the liberalization process and the macroeconomic plan, imports quadrupled from 1985 to 1996, while exports did not even double. As a result, the trade balance dropped from a surplus of US$ 12 billion in 1985 to a deficit of US$ 5.6 billion in 1996. As a consequence of the exchange rate depreciation, the trade balance changed from an annual deficit of US$ 6.5 billion on average during 1996-98 to a surplus of US$ 33.8 billion in 2004, when exports rose from US$ 50 billion on average during 1996-98 to US$ 96.7 billion in 2004. In terms of location issues, the Brazilian manufacturing industry was also affected by these two shocks. Sao Paulo state, Brazil's industry centre reduced its participation in this sector from 52% in 1985 to 43% in 2004. This may indicate that dispersion forces overcome those of agglomeration in Sao Paulo throughout the whole period.

In the Brazilian case, it is feasible to distinguish the effects of contraction and expansion of its economy, which are present at the same time in a trade agreement. After

[48] This is a real devaluation discounted by inflation. The exchange rate, moreover, has not returned to the level before the expansion shock until the end of the period analyzed in this chapter.

liberalization and stabilization plan, domestic production faced fierceness in competition with imports, as they become cheaper in the domestic market. On the other hand, exchange rate devaluation has improved the competitiveness of domestic production since local goods become relatively less expensive than any similar in the international market. As a consequence, domestic firms have expanded their production regardless the target market, either domestic market or exports. In geographical terms, it is important to evaluate how economy reacts in location terms when economy faces more competition (and eventually, it contracts) as well as when it perceives opportunities to increase production (in other words, when firms decide to expand their production).

Summing up, after the liberalization and macroeconomic shock (hereafter referred to as "contraction shock"), less efficient firms facing competition with cheaper imported goods have either since closed down or located somewhere else. Firms importing inputs might have seen an advantage to locate close to the foreign market, but direct effect is stronger, as evidenced by Fally et all (2008), and it overcomes all other effects.[49] On the other hand, after the exchange rate devaluation (henceforth labeled as "expansion shock"), exporting firms may also have reallocated their production plant to regions close to the external market and firms aiming to expand production to attend domestic demand might relocate close to internal market. Therefore, regional disparities in wages may have changed after each shock.

One way to evaluate these shocks is by examining whether transport cost or market potential might have changed its importance to explain regional wage disparities. This chapter contributes to the literature by investigating which shock is stronger to affect regional wage disparities - whether contraction or expansion. Additionally, it is feasible to address which market, internal or external, is more likely to be affected by which shock.

[49] It is relevant to point out that exporting firms using imported inputs can be benefited by reducing cost of imports. However, this is an indirect effect.

Distinguishing which shock and where (internal versus external market) trade issues have a more prominent impact is important for the literature of international trade and regional economics since forces of agglomeration might change after an exogenous shock. Aside from academic interest, these issues are also extremely relevant for policy makers as they become much more aware of the regional consequences of any shock, especially related to trade policy.

The remaining seven sectors of this chapter explore these ideas. Section 3.1 points out the theoretical framework of regional disparities of wages and how they can be influenced by trade shocks. A summary of empirical findings is presented in Section 3.2. Some descriptive analyses are made in Section 3.3. Section 3.4 outlines the econometric specification to test the hypotheses in this chapter, followed by data description in Section 3.5. Empirical results are shown in Section 3.6. Finally, the last section concludes.

3.1. Why do wages differ across regions?

One seminal work in explaining why wages are regionally different is Hicks (1932). His work establishes that wages differ across regions due to two reasons: different costs of living; and amenities. Regions with higher costs of living ought to compensate by offering higher wages. After his work, many different approaches try to explain regional disparities in wages using a wide range of arguments, which are summarized in the next lines.

The heterogeneity of people's skills is the foundation of the Theory of Human Capital, initialized by Becker (1962), to set out why salaries are different and, therefore, also at a regional scale, as laid out by Willis and Orley (1986). According to this theory, regions containing people with more human capital tend to have higher wages. Human capital, indeed, captures wage disparities, but it does not provide us any further

interpretation as to why some regions have better endowments (here represented by human capital); it does not cover labor market demand either by ignoring the existence of firms.

Another part of the literature, called Regional Wage Curve, explains these disparities by different unemployment rates in each region. In other words, the more unemployment one region has, the lower wages are, as explained in Blanchflower and Oswald (1995). This approach takes into consideration not only the demand side but also the supply side in the labor market which eventually culminates into a relationship of unemployment and wages. Although this approach covers the forces in the labor market - a step further from human capital theory - it overlooks important issues. First, why some regions are more prosperous than others, here represented by lower unemployment rate, is not tackled by this literature. Aside that, it also neglects region's characteristics, such as amenities, as explanatory variables to explain regional wage disparities in which they have borne out as relevant aspects, as suggested in Rosen (1979) and Roback (1982).

NEG theoretical models present some further contributions to the existing literature described above. Demand for labor is included in NEG models, since the literature described earlier focuses its attention solely on labor supply. According to NEG literature, firms' location decision provokes higher nominal wages in some areas. The key assumption which drives firms' location is increasing returns to scale. This assumption establishes how agglomeration forces bring about firms' desire to locate close to the market due to consumer demand, as proposed by Krugman (1980) and Krugman (1991), or to forward and backward linkages, as suggested by Venables (1996). Furthering these arguments of agglomeration, Elizondo and Krugman (1995) explains the appearance of metropolises in less-developed countries by stating that a closed economy favors the existence of a domestic industrial center with higher wages. Nevertheless, dispersion forces also exist and competition among firms is considered as the main reason for their location decision.

Overall, firms locating close to the market tend to raise land cost and therefore nominal wages. As a result, regions where agglomeration occurs tend to have higher wages. Consequently, as a region moves away from the market, wages tend to reduce. The main reason for this pattern is the increase of transport cost incurred to ship goods to consumers. Firms facing higher transport cost tend to offer lower wages to their employees. Summing up, firms face a tradeoff between two marginal costs: wages and transport cost.

Although NEG theory findings are in line with previous approaches, they provide a broader view of this phenomenon since supply and demand side for labor is taken into account by acknowledging the existence of firms and their location behavior. The main contribution from NEG literature in relation to regional wages is however how trade may affect these inequalities. Changes in agglomeration and dispersion forces between internal and external markets may influence firms' location decisions by making some regions more attractive than others. Regional wage disparities, as a consequence, might be influenced by these decisions, strengthening or weakening these inequalities.

According to this literature, trade cost is the main exogenous variable which might affect these forces between internal and external market. A reduction of trade cost raises competition with imports for domestic firms and transforms regions close to external markets even more attractive to exporting firms.

3.2. Empirical Evidence of NEG models

As mentioned in the introduction, several papers have addressed whether regional wages can be explained by NEG literature. Two ways are explored: either by transport cost; or by market potential. Seminal papers are Hanson (1996; 1997), in which he has investigated the Mexico case by using transport cost measures. Mexico is a good example to investigate these phenomena because it had a closed economy before signing the North

American Free Trade Agreement (NAFTA) with the USA and Canada in 1985. It is therefore feasible to explore not only whether wages are a decreasing function of distance to markets, but also whether a trade agreement may change these regional inequalities. NAFTA has weakened the importance of Mexico City, the industry center of this country, and increased the agglomeration forces in regions close to the border with the USA. The reduction of trade cost made the external market (USA) more appealing for many Mexican firms. Due to this market effect, firms located close to the border have expanded and others relocated their production plant to these regions. This expansion in production in regions close to the external market may have increased competition in the labor market for employees, therefore raising wages. On the other hand, the fierceness competition with imported goods may lead firms to locate not so close to the border with the USA. Therefore, regional disparities in Mexico might have been changed after this trade shock.

Hanson's results reinforce that transport cost to industry centre or to external market is important to determine regional differences in manufacturing wages, bearing out the assumption of increasing returns to scale from theoretical models. No evidence is found however in either paper while examining whether the trade shock (NAFTA) has influenced regional wage disparities. His explanations on failing to find regional effects include a short period (3 years) after this trade agreement. Nevertheless, the geographical unit used (state level) is also questionable.

Brulhart and Koenig (2006) is another example of the literature on using transport cost to explain regional wages disparities. They evaluate how important access to the European Union market and to the domestic market is for wages disparities in some eastern European countries. Their results do not corroborate NEG predictions, since no robust result related to distance is found.

Another recent approach is using Market Potential instead of distances. Several papers have explored this channel using different measures of market potential and econometric specifications. However, a challenging and substantial difference is how to instrument this endogenous variable. One paper is Mion (2004), in which he evaluates whether market linkages are able to explain the spatial distribution of earnings in Italy. For instruments, he uses the spatial lag of his explanatory variables. Brakman et al (2004) evaluate the same issue by using German district data. For instruments, they use size of districts, population size of each district and population density. Using US county data, Hanson (2005) investigates whether regional demand linkages are associated with wages. Hanson's instruments are related to population in each county. Redding and Venables (2004) use a cross-country data to investigate the same issue: whether demand linkage is able to explain regional wages. As instruments, they use distance to markets represented by New York, Brussels and Tokyo. Head and Mayer (2006) utilize EU regional data to explain how employment and wages are associated to market potential. To tackle the issue of endogeneity, they use not only the distance to Brussels as instruments, but also a distance calculated by the centrality of Europe (named as "EU centrality") as well as "global centrality" measured by the distance to every inhabited place in the world. Although all these papers have different approaches understanding how demand linkages are related to wages, their findings support the idea that market potential is important in explaining those regional disparities, regardless of which instrument is used.

As it regards Brazil, two papers on regional wage inequalities are worth mentioning. Azzoni and Servo (2002) investigate whether regional wage disparities in Brazil may be explained by different human capital endowments. Their paper confirmed this hypothesis by finding human capital as one of the main important features to explain regional wages in Brazil; however, the geographical location also plays an important role. Another similar to

those presented earlier from NEG literature, Fally et al (2008) test whether demand linkages are correlated to Brazilian wage disparities after controlling for individuals' and firms' characteristics. Different from previous NEG empirical studies, the authors also identify which access (market or suppliers access) is more prominent to review these wage disparities. Their result suggest not only that demand linkages are important to explain regional disparities, but also that market access seems to be more relevant than supplier access.

Aside from Hanson (1996; 1997), I am not aware of any other paper investigating whether trade shocks may have influenced regional disparities. This current chapter addresses not only these NEG hypotheses but also whether reduction on trade costs may be relevant to change the strength of agglomeration and dispersion forces.

3.3. Brazilian Regional Wages: A Descriptive Analysis

Some descriptive analyses on relevant variables give us some insights before applying any econometric analysis. The key variable is the ratio of salaries in one region over salaries in the industry center, which is represented by the Sao Paulo metropolitan area in this country.[50] A summary of this ratio is shown at Table 3.1, which splits the information into three periods: before any shock; after contraction shock; after expansion shock.

[50] The regional unit used is microregion. The justification on why this geographical scale is given in Section 3.5.

Table 3.1: Descriptive Analysis of Regional Salaries

Descriptive Summary	Before Shock	After Contraction Shock	After Expansion Shock
Wages in region i / Wages SP	**1985**	**1996-1998**	**1999-2004**
Average	0.44	0.41	0.40
Standard Deviation	0.24	0.23	0.23
Minimum Year-Average	0.10	0.08	0.11
Maximum Year-Average	1.86	1.74	1.76
Number of years available	1	3	6
Number of observations	371	1207	2568
Average of obs. per year	371	402	428

Source: Annual Manufacturing Report (*Pesquisa Industrial Annual – PIA*).

First, the number of observations increases over time, which represents more regions entering in the sample because of the increased number of firms in each locality after the establishment of new firms and/or the enlargement of existent smaller ones.[51] Around 57 microregions (10% of total) have "gained" manufacturing sector plants over the period analyzed, 31 after contraction shock and 26 after expansion shock. This result suggests that some sectors might have decided to locate in other areas, where salaries are lower, due to dispersion forces mentioned earlier.

Looking at these descriptive statistics, it is not clear whether these trade shocks had any impact on how manufacturing wages are regionally distributed. There is a small reduction of the average over the period analyzed. Additionally, the maximum value declines throughout the whole period analyzed as well. Other descriptive statistics, such as minimum and standard deviation, present no substantial differences. Overall, it is difficult to get insights from these non spatial measures, which lead us towards a more regional investigation into this variable later on.

[51] According to the source, all manufacturing plants over 30 employees are included in the data. Therefore, the inclusion of more regions show that either firms have become larger over the years in regions with no record previously or new firms have established themselves in those regions. Firms having less than 30 workers do not enter in the dataset.

Table 3.2 shows a summary of distances, either to the industry center, or to the foreign market, which is represented by the nearest port in this case.[52] Regions are closer, on average, to the external market (nearest port) than to the industry center (Sao Paulo). This can be partially explained by how these distances are calculated, since port distance is created by nearest, but there is only one industry center. Nevertheless, most of the microregions are located closer to the coast, which sets out another part.

Table 3.2: Descriptive Analysis of Distance to Markets

Distance (in km)	SP	Port	After Suape	After Sepetiba	After Pecem
Average	1,291	427	398	396	393
Standard Deviation	832	295	303	303	304
Minimum	14	22	22	22	22
Maximum	3,317	1,418	1,418	1,418	1,418
Regions changed after infrastructure improvements (137)			64	36	37

Regarding transport cost to external markets, during this period three ports were inaugurated and/or hugely expanded aiming to improve foreign market access, one in the Southeast (Sepetiba in 1998) and two in the Northeast (Suape in 1998 and Pecem in 2002). These modifications in the port system make the distance to the external market vary over time. Since these port improvements, the average minimum distance to the external market has reduced consistently (from 427 km to 393 km). Moreover, 137 microregions have become closer to the nearest port since these improvements.[53]

In order to have an overview of the geographical change, some maps are shown to visualize how these trade shocks have affected regional wages in Brazil. Two exercises are made. First, an average of the ratio of salaries from 1996 to 1998 is compared to wages in

[52] Most of the Brazilian International trade is done by maritime transport. More details are available in Section 3.5.
[53] Changes at the market potential measure are only marginally different; therefore, it is not shown here.

1985 to investigate the contraction shock. If changes were greater than a certain limit, then these particular regions have become more similar (or less, if it is negative) to the industry center.

Three different percentage limits are used for this purpose: 1%, 5% and 10%.[54] Nevertheless, interpretations remained identical regardless the percentage limit used. For simplicity, maps showed in this section use only 5% as a limit.[55] In all maps presented, regions are classified as "A" if they have become more similar to Sao Paulo, "B" if they have not shown a significant change and "C" if they have become less similar to the industry center. There are some regions where information is not available either before and/or after, which are denominated as "D". Figure 3.1 show how manufacturing wages at a mesoregion scale have changed after the first liberalization shock.[56]

Numerically, almost the same number of regions have become more (48) and less (53) similar to the industry center. However, proportionally regions in the South have benefited more than any other part of Brazil,[57] especially on the Parana and Santa Catarina coast, which includes some other important manufacturing regions in Brazil.[58] One explanation as to why the South has been more affected is Mercosul, a regional trade agreement signed by Brazil, Argentina, Uruguay and Paraguay, which started at the beginning of the 1990s. Another geographical pattern is that regions on the coast seem to have benefited the most, while regions in the countryside had the opposite experience. Only 21 regions do not present any substantial difference after the first trade shock. These regions do not show any particular geographical pattern. In summary, these results suggest

[54] On average, regions have closed the gap by 3% overall. Considering the 95% confidence interval of the difference, the lower bound limit is -4% and the upper 11% in the contraction shock, while 0% and 6% in the expansion.
[55] Maps using 1% and 10% as limits are available in Appendix 3.IV. As previously stated, they show similar geographical interpretations.
[56] For a matter of better visualization, mesoregion scale is used instead.
[57] One explanation as to why the South has been more affected is Mercosul, which started at the beginning of the 1990s.
[58] These are the Curitiba metropolitan region in Parana and the Itajai Valle in Santa Catarina.

that regions on the coast, and particularly in the south, are those which show some catch up to the industry center, while countryside regions not.

Figure 3.1: Map after Contraction Shock

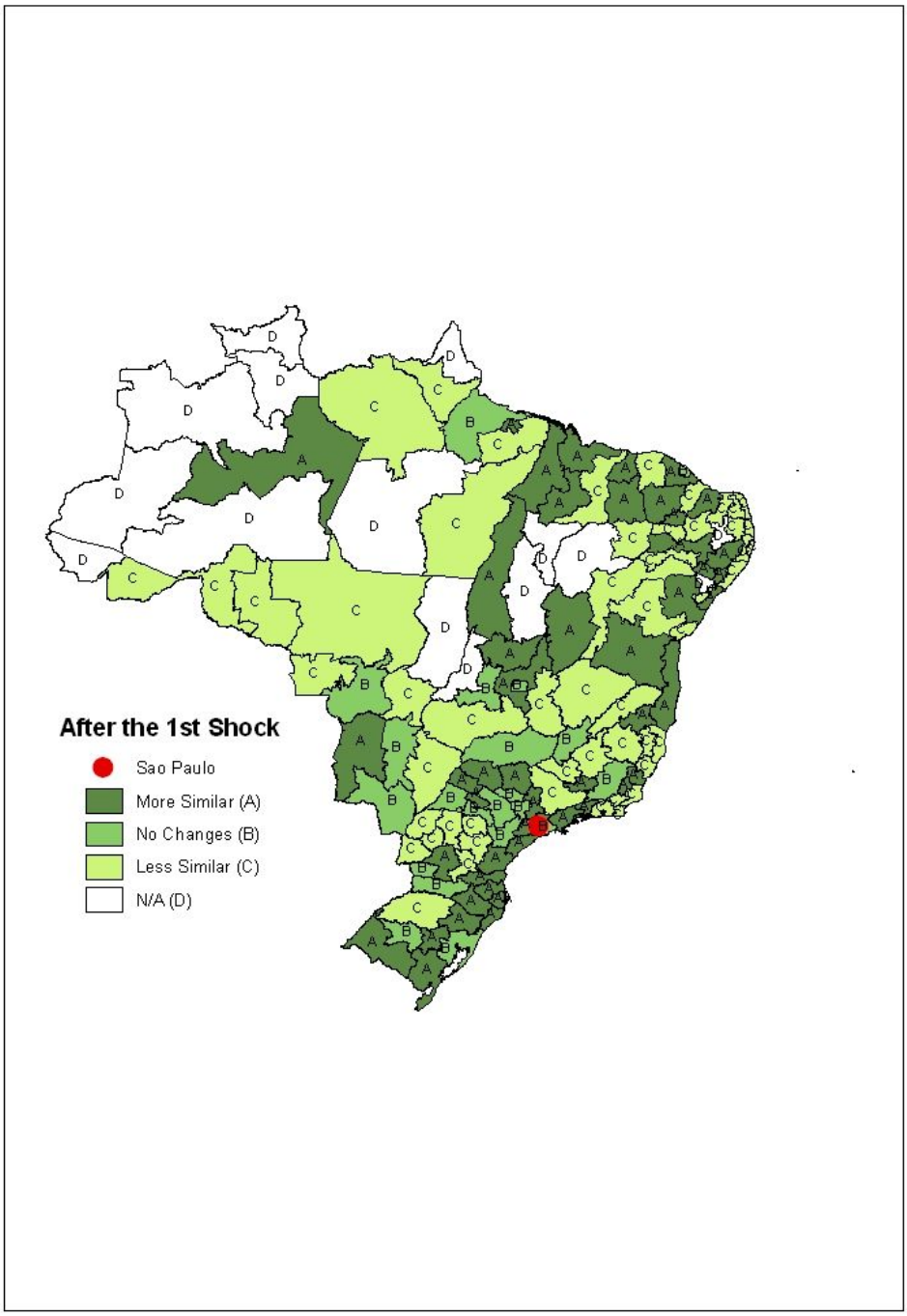
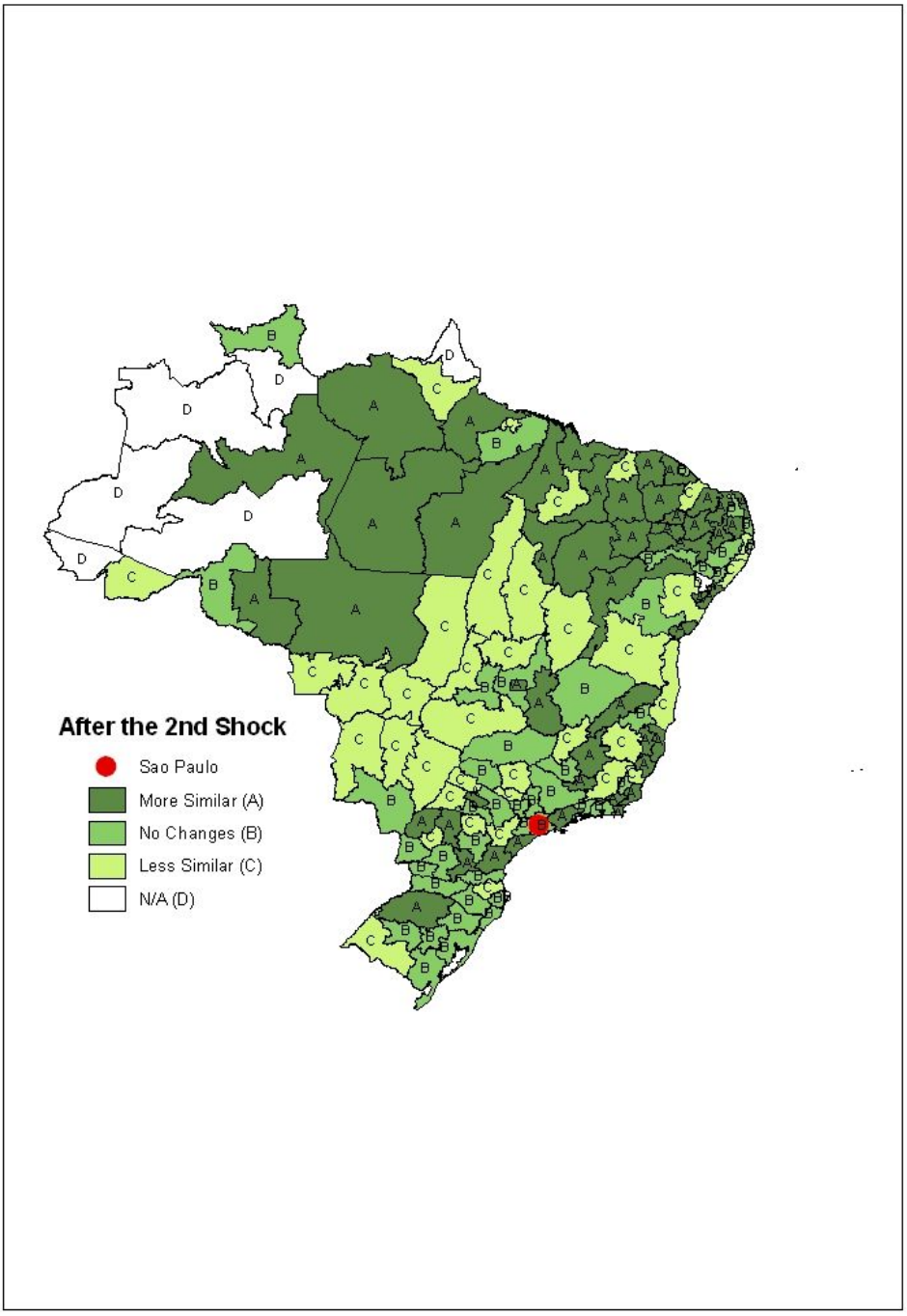


Figure 3.2 shows the same analysis but for the expansion shock: exchange rate devaluation. For this investigation, average of salaries from 1999 to 2004 is compared to salaries from 1996 to 1998. The group classification remains the same as before.

Figure 3.2: Map after Expansion Shock



After the exchange rate devaluation, more regions have become more similar to the industry center (49) compared to those which have become less similar (37). Another different feature from the liberalization process is a greater number of regions (44, more than double of 21 after the contraction shock) which have not shown any substantive difference.[59] These figures suggest, in general terms, that the contraction shock seems to have impacted more regions (101) than the second (86), but leading them to a much more unequal distribution of wages across the Brazilian landscape. Nevertheless, the negative impact appears to happen much more in the liberalization process than in the exchange rate devaluation.

Proportionally, North regions seem to have caught up with the industry center compared to other parts of Brazil after the expansion shock. Some metropolitan areas have also benefited from both shocks, like the Curitiba metropolitan area in the South, while others only after the expansion shock, such as the Salvador metropolitan area in the North. Nevertheless, other regions from the southeast have also reduced the gap to Sao Paulo. One common feature of the liberalization process is the fact that coastal regions seem to benefit more than countryside regions, where most places have increased the gap with Sao Paulo.

These outcomes present some insights. Coastal regions, regardless of whether after the contraction or expansion shock, tend to have become much more similar to the industry center than regions in the countryside. This suggests port regions, or those close to, are the most affected by any trade shock, since most of Brazilian international trade is made through maritime transport. Differences are however distinguishable between the North and South, while southern regions closed the gap after liberalization; northern regions did so after the exchange rate devaluation.

[59] Similar patterns occur if 10% is used instead. Using this threshold, 21 become 44 after the contraction shock and 44 becomes 74 after the expansion shock.

3.4. Empirical Strategy: Measuring how Geography Matters and Whether Trade Shocks Impact Regional Wage Disparities

As mentioned previously, Brazil experienced some economic shocks: a liberalization process in the early nineties, a macroeconomic stabilization plan in 1994 and exchange rate devaluation at the end of 1990s. An empirical strategy ought to have the capacity to explore not only the NEG literature by showing how demand linkages are able to explain regional wages inequalities, but also whether these trade shocks have impacted these disparities and, if so answered affirmatively, which was more effective.

Testing whether demand linkages can explain regional wage is done by including transport cost and market potential as explanatory variables, but not jointly since they are almost the same measure. The initial approach is estimating with transport cost followed by using market potential. One crucial concern using market potential resides in the fact that it is an endogenous variable and needs to be instrumented. The key issue of this empirical strategy is to detect whether trade shocks have affected or not these disparities. This is done by trying to evaluate whether the importance of the variable investigated has been reinforced or not over the years after the shock. In other words, I test whether the slope of transport cost or market potential has changed over time.

As discussed earlier, differences in regional manufacturing wages may be explained by transport costs. Equation 3.1 presents the econometric specification to capture how it explains these disparities.

(3.1)

$$\ln(W_{it}/W_{ct}) = \beta_0 + \beta_1 \ln(IC_{it}) + \beta_2 \ln(PORT_{it}) + \delta_t\theta \ln(IC_{it}) + \delta_t\lambda \ln(PORT_{it}) + \gamma_t\varphi \ln(IC_{it}) + \gamma_t\rho \ln(PORT_{it}) + \alpha_k \ln(Control_{kit}) + \varepsilon_{it}$$

where W_{it} is the average nominal wage per worker for region *i* at time *t*; W_{ct} is the average nominal wage per worker from the industry center in Brazil, the Sao Paulo Metropolitan

Area[60], at time *t*; IC_{it} is the unit of transport cost from region *i* to industry center at time *t*; $PORT_{it}$ is the unit of transport cost from region *i* to the closest port at time *t*; $Control_{kit}$ is the control variable *k* from region *i* at time *t*; δ_t is a dummy variable which takes a value one if year *t* falls after the contraction shock; γ_t is a dummy variable which takes a value one if year *t* falls after the exchange rate devaluation; ε_{it} is the error term, which is discussed later; the remaining terms are parameters to be estimated.

According to the theory, a negative estimated value of β_1 and β_2 is expected since an increase in transport costs reduces the value of (W_{it} / W_{ct}), which means that the salary of a region far from the industry center or from the nearest port becomes lower relative to the market. In summary, parameters β_1 and β_2 test the following hypothesis:

a) If β_1 and/or β_2 are significantly negative, transport costs matters for difference in regional wage;

b) If β_1 and/or β_2 are not significant, transport costs are irrelevant to understand regional wages disparities.

If β_1 and/or β_2 change after a trade shock, then liberalization and/or exchange rate devaluation have impacted regionally the Brazilian economy. In order to test this, it is possible to check whether Equation 3.1 is stable over time, which is equivalent to test that θ and λ are equal to zero for the contraction shock and φ and ρ for the expansion shock.

To examine which channel of the dispersion force (increase in competition) is higher in the Brazilian case, this could be seen by comparing the coefficients on internal market after the contraction shock (θ) with the other after the expansion shock (φ) and coefficients on external market after the contraction shock (λ) with the similar after the

[60] This is the Sao Paulo microregion and not the state or the city of Sao Paulo.

expansion shock (ρ). If $|\theta| > |\varphi|$ and $|\lambda| > |\rho|$, then competition with a new product (imported goods) has influenced the reduction of disparities of wages more than competition in the labor market.

Not only demand linkages design regional disparities as discussed previously; regions containing people with higher skills tend to have greater wages, as stated in human capital theory. Three measures are used for controlling this issue: years of schooling, human capital and productivity. The second uses the standard measure of human capital[61], while the first is derived by average number of years spent on schooling. Productivity, measured by value added per employee, is also used to control for regions' labor quality. In order to provide some exogeneity of this variable, human capital measure is lagged in time. However, people may argue that this variable is still endogenous, since this sort of variable changes slowly over time. It is then relevant to find a suitable instrument for this endogenous variable. First, human capital is measured by a combination of education and working experience. Therefore, education is just a fraction of human capital and sometimes not sufficient to get a job, since some vacancies ask for either working experience or specific technical skills. Moreover, it seems that illiteracy rates in any region may only affect wages through human capital and not through any other measure. If that is the true, this feature satisfies the exclusion restriction of instruments. Additionally, Brazil has a substantial proportion of illiterate people in its population, despite reducing it since the 1950s where approximately half of Brazilians were illiterate. According to the last Brazilian Census in 2000, around 13,6% of inhabitants beyond 15 years old is still unable to read or write. Moreover, those figures vary substantially on geographical terms, where some regions have already eradicated illiteracy, while others present illiteracy rates much higher than the National level. Those figures represent how illiteracy is important in the

[61] This includes schooling and work experience. More details can be found at www.ipeadata.gov.br.

Brazilian society as well as how they might be relevant in regional aspects. Aside those arguments, the exogenous nature of illiteracy rate is strengthened by the Brazilian context. There is a trade-off experienced by some low income Brazilian families: either enrolling their children at school; or asking them to contribute to the family budget during the time they would be at school. Those families represent a great share of Brazilian illiterate inhabitants. Therefore, the reverse causality of income over illiteracy may not apply, since some (yet not all) of those families may increase their income by not allowing their children to learn how to read and write. To substantiate this, regional wages are regressed on human capital measure, and the latter is instrumented by the percentage of illiterate people from each region, since illiteracy may impact regional wages only via human capital measure, therefore satisfying exclusion restriction needed for an instrument.

Regions are not homogeneous in terms of industry share, while in some the economy relies more on the service sector; in others manufacturing may play an important role. Therefore, the relevance of the manufacturing industry in each locality ought to be taken into account as some regions might have higher wages due to some manufacturing firms offering better remunerations. Three measures are utilized for this purpose: the share of manufacturing in regions' GDP; the share of manufacturing employment; last, but not least, the percentage of number of manufacturing firms. The first uses the ratio of the value added in the manufacturing sector over regional GDP, while the second shows the share of manufacturing employment among the total urban population. The last one is calculated by showing the percentage of firms in each region over the total Brazilian manufacturing sector.[62]

Brazil has experienced a dispute between states to attract manufacturing plants in the 90's, referred to as the "Fiscal War", where states offered exemption of taxes and other

[62] Since the dependent variable is related to the industry centre, all these controls are also measured in relation to Sao Paulo for estimation purpose.

subsidies to achieve this goal as evidenced by Rodriguez-Pose and Arbix (2001). Therefore, it is important to control for this government intervention. Two channels are explored. One uses the percentage of expenditure in regional development and manufacturing industry over the total budget of each state. The other assumes there is an exogenous state effect to attract manufacturing plants. Then state fixed effects are utilized to capture this policy intervention.

As discussed above, infrastructure improvements to external market were made by government spending in three ports and these ought to be included in this estimation process. Two ports in the North, Pecem (CE) and Suape (PE), were basically inaugurated for international trade during the period investigated in this chapter, since figures on import and export appears only after a certain year[63]. Another port in the South, Sepetiba (RJ), was extremely expanded from 1998 onwards. The volume of international trade in this port doubled from 1997 to 1998. It also increased substantially every year after that. By 2004, the amount of imports plus exports has been multiplying by 12 compared to 1997 figures. Thus, the distance to the nearest port is measured by including these infrastructure changes, where these three ports were included as an option only after major changes were made by government expenditures.

Panel data approach using regions as individual unit is essential for this study, since it controls for any time invariant regions' characteristics. Fixed effects ought to be used for this purpose, but they present a different pattern in comparison to Hanson (1997). Hanson mentioned three exogenous amenities which are captured by fixed effect in the Mexican case: (i) exogenous natural-resource supplies; (ii) exogenous levels of amenities; and (iii) location bias in government spending or tax policies. Although the first two are also exogenous in the Brazilian case, the last one can not be considered fixed for this country, as

[63] Pecem (CE) after 2003 and Suape (PE) after 1998.

stated in the last two paragraphs. The error term has a specific form based on the fixed-effect approach displayed in Equation 3.2.

$$\varepsilon_{it} = c_i + v_t + \eta_{it} \tag{3.2}$$

where c_i is the fixed effect for region i, v_t is the fixed effect for year t, and η_{it} is an i.i.d. term with mean zero and finite variance σ^2.

Estimation by fixed-effects presents a problem. The transport costs used in Equation 3.1 do not vary over time, but rather within regions, since it is the distance from a region to the industry center or to a port. The distance to the external market is solved by the introduction of improvements in infrastructure (ports' inauguration), as stated previously. The remaining problem is distance to industry center. First-differencing the data would eliminate the distance variables from the regression. One way to overcome this issue is by multiplying the distance to the industry center importance in the economy over the years. The share of the industry center in the Brazilian economy GDP is used as a measure of "importance". Since industry center relevance is not static over time, the independent variable becomes time-variant.[64] Fixed-effects approach becomes feasible after this modification.[65]

As mentioned earlier, another approach to detect demand linkages is how much market potential can explain regional wage disparities. In order to probe these transport cost results, estimation using measures of market potential may provide further evidence. First, it is important to define what internal and external market potential is and how they are calculated. For the former, the GDP from each microregion is used to construct the internal market potential and is calculated by using formula 3.3.

[64] Actually it seems to have lost share, as suggested in the introduction.
[65] Random effects are also used in order to use state fixed effect to capture the government incentives.

$$IM_{it} = \sum_{j} \left(GDP_{jt} \Big/ DIST_{ij} \right) \tag{3.3}$$

where IM_{it} is the internal market potential of region i at time t; GDP_{jt} is the GDP of microregion j at time t; $DIST_{ij}$ is the distance between regions i and j;

The external market potential is created by using data from international trade of Brazilian ports.[66] The export volumes added to import value of each port give us the external market of each port region. Additionally, foreign market potential is measured by a similar expression for internal market potential by using the distances between regions and ports. Formula 4 shows how it is calculated.

$$EM_{it} = \sum_{j} \left[\left(X_{jt} + M_{jt} \right) \Big/ DIST_{ij} \right] \tag{3.4}$$

where EM_{it} is the external market potential of region i at time t; X_{jt} is the total exports of port j at time t; M_{jt} is the total imports of port j at time t; $DIST_{ij}$ is the distance of region i and port j;

Instead of utilizing distance to the industry center or ports, market potential measures (internal and external) of each region is employed as an alternative in an econometric specification. Equation 3.5 shows the new econometric equation.

$$\begin{aligned} \ln(W_{it}/W_{ct}) &= \beta_0 + \beta_1 \ln(IM_{it}) + \beta_2 \ln(EM_{it}) + \delta_t \theta \ln(IM_{it}) + \delta_t \lambda \ln(EM_{it}) \\ &+ \gamma_t \varphi \ln(IM_{it}) + \gamma_t \rho \ln(EM_{it}) + \alpha_k \ln(Control_{kit}) + \varepsilon_{it} \end{aligned} \tag{3.5}$$

where W_{it} is the average nominal wage per worker for region i at time t; W_{ct} is the average nominal wage per worker from the industry center in Brazil, the Sao Paulo Metropolitan Area, at time t; IM_{it} is the internal market potential of region i at time t, as defined before; EM_{it} is the external market potential of region i at time t, as defined before; $Control_{kit}$ is the

[66] The next section provides an explanation on why using only port.

control variable *k* from region *i* at time *t*; δ_t is a dummy variable which takes a value one if year *t* falls after contraction shock; γ_t is a dummy variable which takes a value one if year *t* falls after exchange rate devaluation; ε_{it} is the error term, which is discussed later; the remaining terms are parameters to be estimated.

Regarding the issues discussed in the previous econometric specification, most of them remained practically the same, but now fixed effects can be used because market potential is not constant over time. Another feature, however, should be examined here with care. Market potential, regardless of whether internal or external, is endogenous in this econometric specification. Endogeneity of regressors leads to biased estimates of the parameters and instruments are therefore required. Instruments should be able to affect dependent variable only through endogenous variable in order to satisfy their two main conditions: correlated to the endogenous variable; and orthogonal to the error term.

Brazilian government provides long-term loans through its development bank, named BNDES. There are different reasons defining which projects should be financially supported by BNDES loans, especially in geographical terms. Two of the most important reasons are either economic or police. Some greenfield projects, for example, need to be implemented in regions where economies of scale are crucial for project's profitability. Therefore, such a project should be implemented in one of the most developed regions within Brazilian territory; otherwise this project is not financially viable. On those projects, rich regions are benefited by those financial supports and economic reason prevails. On the other hand, some projects are designed to promote development in poor regions. One example might be investments in infrastructure in less developed regions in which the goal is to boost local economy. For those projects, loans are driven for police reasons and they favour poor regions within the Brazilian territory.

Summing up, BNDES loans have therefore not a single pattern on where disbursements should be regionally allocated, since they depend on distinctive projects received by the bank. Moreover, rich and poor regions might be benefited idiosyncratically through years. Just to exemplify, the richest state in Brazil (Sao Paulo) has received only 0.8% of disbursements over its GDP between 1991 and 1999, while one poor state (Sergipe) the same ratio was 2.0% over the same period. On the other hand, during the same period, another poor state (Piaui) has received only 0.5% of disbursements over its GDP, while another rich state (Santa Catarina) shows 1.7% of the same ratio. Looking at those figures, it is safe to conclude that because BNDES has different objectives for its loans, they do not have, for example, any economic or police pattern. In this respect, it seems that BNDES disbursements would appear to be plausible instruments, but some further precautions might be appropriate in order to tackle each endogenous variable, internal and external market potentials.

BNDES has different types of loans, some for exports and others aiming the domestic market. Since there are two endogenous market potential, one internal, another external, loans are divided according to each market. All export oriented loans are used to instrument external market potential; others types of loans serve as instrument for internal market potential.

In order to affect dependent variable only through endogenous variable, some additional precautions are taken. First, disbursements are lagged in time in order to provide some time for projects to be fully implemented. As exports loans last one year, disbursements are lagged one year. On the other hand, disbursements for internal market last for 5 years on average, then they are lagged for the same period. This time lag helps to avoid any excess in demand for labour due to exogenous increase of investments, since labour market might clear after the implementation of the project due to the attraction of

migrants, for example. Second, all values are at state level, not at microregion level as wages and market potentials are. Due to this more aggregated regional unit, any government intervention affects all different microregions in the same state evenly. Then, any microregion in each state benefits from government support, especially those closer to where investments are made. Third, BNDES disbursements are divided by each regions' GDP or Exports, which makes them even more exogenous and probably only related to market potential and not to regional wages.

Gathering those information, these bank's disbursements over each regions' GDP or exports might be feasible instruments, since BNDES loans may only impact wages through GDP, which is the basis of market potential measure. If that is true, those government interventions might be able to satisfy the exclusion restriction, one of the conditions that an instrument should be respecting.

It is not possible to control for all possible variables that might be correlated with BNDES disbursements and regional wages. Furthermore, this empirical approach might capture the effect of BNDES disbursements on regional wages, but working through other different channels. In order to tackle these problems, other variables used in previous papers are performed as additional instruments for the same purpose: to instrument market potentials. Brakman et al (2004) and Hanson (2005) use population as instrument for market potential. Therefore, the share of urban population lagged in time represents a size measure which might be related to nominal wages only through market potential in this research, since manufacturing jobs are generally available in urban areas. Distance measures are exogenous characteristics enabling them to become an instrument, but they seem to be related to wages. If that is true, distances to markets are not the ideal instruments, since they may fail the exclusion restriction. However, distance to economic centers is used as instruments in previous studies, such as Head and Mayer (2006) and

Redding and Venables (2004). Indeed, Head and Mayer (2006) tried two other measures of distance which are considered more exogenous[67], but no real difference is found in their paper. Hence, these additional measures appear to be the most common instruments used in the literature for market potential measures. Moreover, expanding the set of instruments may lead to more efficient estimation and enable this research to perform over-identification tests.

Summing up, regional wages are regressed on internal and external market potential, and those endogenous variables are instrumented by BNDES disbursements over GDP or Exports (depending on which market potential) jointly with population urban share lagged in time and distances to internal and external market of each region. The main reason for choosing those variables is that they might affect regional wages only through market potential, satisfying the exclusion restriction which is one crucial condition for their validation. Nevertheless, those instruments should meet those criteria by statistic tests, which are supportive, not conclusive.

3.5. Data Description

One major issue is to define which geographical unit could better capture these effects. Even though Hanson (1997) analyzed this effect using Mexican states data, he argues that a more disaggregated data, for example split by cities, could be more suitable for this investigation. His justification to use state-level data rests on the fact that manufacturing employment is concentrated in one single city at each Mexican state. Although this could be the case of Mexico, this is not the Brazilian case. There are many important cities within states which cannot be neglected by pooling all of them together in one single unit (states). One aspect to be considered in selecting a geographical scale is whether a political division represents an economic one. Regional disparities within any

[67] These distances are named as "EU centrality" and "global centrality".

larger scale politically established cannot be captured by the data, for example at a state level. Therefore, movements to poor regions in the same geographical unit are not able to be captured by data in such geographical scale. On the other hand, geographical scale ought to be economically meaningful.

Apart from the political division into states and cities, the Statistic and Geography Brazilian Institute (*Instituto Brasileiro de Geografia e Estatística* – IBGE, henceforth) has two other geographical classifications: mesoregion; and microregion.[68] The former divides the territory into 137 parts and the latter into 558. City-level data are not the most appropriate for this period for two reasons. First, more than 1,000 districts were transformed into cities by emancipating themselves in the nineties. Another issue resides in the fact that municipality-level splits functional areas, which ought to be treated jointly. Considering these issues, microregion division seems to be the most suitable spatial scale for the investigation proposed at this chapter.

The main data are gathered from IBGE. Two types of publication are used: Annual Industry Research (*Pesquisa Industrial Anual* – PIA) from 1996 until 2004; and Industrial Census from 1985. The former has annual information from all establishments of over 30 employees and the latter has information from all manufacturing firms; however, only firms above the threshold of PIA are used for analysis in order to keep the same sample characteristics from both publications.

The investigated variable is wages, which is measured by the total remuneration divided by the number of employees in each microregion. In other words, average wage per worker at each location. Transport cost is calculated by the distance between the main city of a microregion and the industry centre (Sao Paulo) and the external market.

[68] These two classifications use the social and economical linkages to evaluate which cities are more connected, but they respect the political division (states and cities). Thus, they are nested classifications between states and cities.

The Mexican case has a particular geographical location, since it shares a common border with its most important trade partner (the USA). This facilitates any study of trade geographical impact in this particular country. Brazil international trade is evenly distributed with all parts of the globe. The highest percentage does not exceed 30% over the whole period. The UK has a similar pattern; moreover, a great part of its international trade is by maritime transport. Overman and Winters (2006) tackle this problem by using the distance to the ports as a proxy of distance to the international market in the UK case. Brazil has a large number of ports along its extensive coast, but the main 14 Brazilian ports represent more than 60% of the total international trade.[69] These ports are selected according to Goebel (2002) and Lacerda (2004) based on their historical data and capacity. Only the minimum distance to ports is used in this chapter and average distance are neglected, since the former represents more adequately foreign market access than the latter. Consider one simple example to understand why: imagine a region as a straight line with two ports in each extreme. The average distance to foreign market does not have any difference for all random point selected at this region, since the mean distance to a port is the same for all of them. It is clear, however, that the port regions (the extremes) have better access to the external market than any inside region. Minimum distance to a port is thus the best proxy for external market access. The distance is given by the Great Circle Formula using the latitude and longitude of each location (city and/or port).

As stated previously, demand linkage can be explained by market potential. Internal market potential is calculated by using GDP measures created by IBGE and Applied Economic Research Institute (*Instituto de Pesquisa Economica Aplicada* - IPEA). External market potential is created by utilizing the total of trade (imports plus exports) at all ports

[69] The ports are four in the Southeast (Santos-SP, Rio de Janeiro-RJ, Sepetiba-RJ and Vitória-ES), four in the South (Paranaguá-PR, São Francisco do Sul-SC, Itajaí-SC and Rio Grande-RS), four in the Northeast (Salvador-BA, Fortaleza-CE, Suape-PE and Pecém-CE) and two in the North (Belém-PA and Manaus-AM).

mentioned before, sourced by the Trade Secretary. Distance to construct these measures is calculated by Great Circle Formula.

Regarding human capital measures, different sources are used. Human capital, in a strict sense, is calculated by IPEA using education and work experience.[70] Education is derived by average number of years spent in schooling for any citizen beyond 25 years of age presented in the Brazilian Census of IBGE, which is the same source for illiteracy of adults over 25 years of age. Productivity of each region is measured by dividing value added by number of employees, sourced by PIA.

For controlling manufacturing importance, most of the measures originated from PIA and the remaining from IPEA. The source of states' expenditure, which is utilized to capture subsidies for manufacturing, is from the National Treasury.

3.6. Results

As a matter of simplicity, results shown in this section take into consideration only one measure of labour quality (human capital) and one measure of manufacturing importance at each location (percentage of manufacturing's value added in each region GDP).[71] Both measures of government regional policies, percentage of subsidies and state dummies, are presented here.[72] Since human capital is endogenous, the percentage of illiterate people is used as it may affect regional wages only through human capital, as explained previously.[73] Period effects are included to capture any time change which is not related to the phenomenon investigated. Last, but not least, errors are robustly estimated.

[70] More details see www.ipeadata.gov.br.
[71] Outcomes with other different measures, such as productivity and years of schooling for labour quality and share of employment and number of manufacturing firms for manufacturing regions importance present similar results to what is shown in this section. These further results follow attached in the Appendix 3.II.
[72] Parameters estimated for state dummies are not shown. Whenever subsidies parameters estimated are not included in any column, it means that state dummies are used instead. However, state dummies cannot be used with fixed-effects for obvious reasons.
[73] As dependent variable is measured in relation to Sao Paulo; then, human capital measures used here are also divided by Sao Paulo figures.

Table 3.3 provides the first results using equation 3.1 and it is structured as follows. The first six columns (i to vi) present results with time-invariant distance to industry center, while the last four columns (vii to x) show using the distance to the industry center weighted by its importance. The first column of each method (i, iv and ix) presents results with no trade effect. Columns ii, v, vii and x use state government expenditure to control for subsidies.[74] Those remaining (iii, vi and viii) show results using state dummies instead.[75]

Table 3.3: Distance as Explanatory Variable– Equation 3.1

Panel A: Regression Results										
Dependent Variable Wi,t / Wsp,t	IV			IV Random Effects					IV Fixed Effects	
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Dist. SP	-0.11	-0.11	-0.05	-0.13	-0.12	-0.08	-0.14	-0.11	0.02	-0.29
	(0.013)***	(0.036)***	(0.022)**	(0.024)***	(0.026)***	(0.034)**	(0.025)***	(0.038)**	(0.06)	(0.055)***
SP Cont. Shock		0.00	-0.03		-0.01	-0.02	-0.01	-0.02		0.02
		(0.04)	(0.02)		(0.01)	(0.01)	(0.01)	(0.01)		(0.011)**
SP Exp. Shock		-0.01	-0.01		0.00	0.00	-0.01	-0.01		0.02
		(0.04)	(0.02)		(0.01)	(0.01)	(0.01)	(0.01)		(0.011)*
Dist. Port	0.04	0.07	0.02	0.02	0.06	0.04	0.06	0.04	0.02	0.05
	(0.009)***	(0.023)***	(0.02)	(0.011)**	(0.014)***	(0.014)***	(0.014)***	(0.014)***	(0.013)*	(0.015)***
Port Cont. Shock		-0.03	-0.04		-0.04	-0.04	-0.04	-0.04		-0.01
		(0.03)	(0.018)**		(0.011)***	(0.011)***	(0.011)***	(0.011)***		(0,01)
Port Exp. Shock		-0.03	-0.04		-0.05	-0.05	-0.05	-0.05		-0,03
		-0,02	(0.017)***		(0.010)***	(0.010)***	(0.010)***	(0.010)***		(0.008)***
Human Capital	0,37	0,37	0,17	0,29	0,29	0,17	0,28	0,17	0,11	0,14
	(0.030)***	(0.030)***	(0.013)***	(0.048)***	(0.045)***	(0.027)***	(0.043)***	(0.026)***	-0,15	-0,10
Subsidies	-0,05	-0,05		-0,02	-0,02		-0,02		-0,01	-0,01
	(0.007)***	(0.007)***		(0.006)***	(0.006)**		(0.006)***		-0,01	(0.006)*
Manuf. VA / GDP	0,09	0,09	0,13	0,08	0,08	0,09	0,08	0,09	0,09	0,08
	(0.008)***	(0.008)***	(0.007)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***
Period Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Effect	No	No	Yes	No	No	Yes	No	Yes	No	No
Obs.	4.145	4.145	4.145	4.145	4.145	4.145	4.145	4.145	4.145	4.145
R-squared	0,23	0,23	0,52	0,38	0,38	0,51	0,39	0,51	0,30	0,43
Panel B: First Stage Results										
Instrument	Illiteracy	illiteracy	illiteracy	illiteracy	illiteracy	illiteracy	illiteracy	illiteracy	illiteracy	illiteracy
FS R-sq.	0.09	0.09	0.27	0.19	0.19	0.19	0.19	0.19	0.16	0.17
FS F-stat	388	388	1,520	35	35	35	35	35	225	278

(a) Robust standard errors in parentheses; (b) * significant at 10%; ** significant at 5%; *** significant at 1%.

[74] Hausman tests are performed between vii and x. Results reject the hypothesis that both estimation provide similar parameters estimated. Therefore, fixed effects present more robust results.

[75] In all these columns, human capital measures are instrumented by illiteracy rate.

Regarding rows, panel A presents estimation results, while panel B show First Stage information, where the first row present which instrument is used, the second reports R-squared and the last, F-statistic. It is not possible to show any over-identification test, since IV approach is just identified.[76]

It is important to observe some facts before interpreting the results. Illiteracy seems to explain substantially human capital measure, since R-squared is over 0.15 in most first stage results. F-statistic also shows that this variable is statistically significant to explain human capital. Therefore, there is some evidence that illiteracy might be a valid instrument.

Regarding interpreting controls' outcomes, human capital plays an important role in explaining differences in wages, since most estimation results show the expected sign (positive) significant.[77]The importance of manufacturing in each microregion also presents a positive and significant sign, which strengthens our results by controlling how relevant this sector is at each location.

The other control variable, subsidies, does not show consistent outcomes. First, it seems that it has a negative impact, contradictory to what is expected. Regions with higher subsidies to industry in their budget, thus, have lower wages. This finding is consistent with Sousa (2002), which also found that states with higher subsidies didn't attract more manufacturing production, but contrary to Volpe (2004). When states dummies are included, the majority does not present significant results, showing that any time invariant state characteristic is not important to explain differences in regional manufacturing salaries. This result in subsidies shows that more work should be done to address the

[76]Over-identification test are presented when using Market Potential later in this chapter.
[77] Outcomes with education and productivity present similar results (see Appendix 3.II), as well as not instrumenting human capital (Appendix 3.I).

question of to what extent the Fiscal War has really played a role in the location of manufacturing sector in Brazil.[78]

As it regards transport cost, it is clear that distance to the industry center plays a relevant role in explaining regional wages in Brazil, since it has the expected sign showing that transport cost really matters. The transport cost measure to external market presents however the opposite. The farther a region is from the external market, the higher wages are, since it presents a positive significant sign. Internal market thus appears to determine the regional distribution of salaries in Brazil, but external market seems to explain the other way around.

By analyzing how trade shocks have impacted these disparities, it is possible to point to an increased importance of external market and the opposite for the internal market, as shown in lines Sao Paulo or Port after contraction and expansion shocks. However, robustness differs. Distance to the industry center becomes less important after both trade shocks, only in the FE approach. On the other hand, transport cost to ports shows more robust results regardless which method is used. However, distance to external market is only affected after expansion shock when FE approach is used. Despite those differences, these findings suggest that in a closed economy (1985 in this example); transport cost to internal market shapes how manufacturing wages are regionally distributed. After opening to trade, distance to external market tends to increase its importance in explaining regional wage inequalities compared to domestic market since distance to internal market after either shock has a significant positive sign in the FE approach and distance to external market has a significant negative one in most cases. These effects allow me to set out why

[78] Another part of Brazil where manufacturing receive government intervention is Manaus Metropolitan Area as designed by the Manaus Free Trade Zone. Creating a dummy for this metropolitan region captures this effect, but it cannot be used in fixed effect approach. Nevertheless, results using RE and OLS do not show any difference.

regions in the South (after liberalization) and in the North (after exchange rate devaluation) have closed the gap to the industry center, as indicated in the map from Section 3.3.

Previously, Hanson (1997) found no evidence of NAFTA affecting Mexican regional wage disparities, which is explained in the paper by the short period considered after the trade agreement. It is important to notice another issue: geographical unit, which is at a state level in Hanson (1997). A more detailed geographical unit and time lag after a trade shock illuminate these effects on regional wages disparities as outcomes of this chapter highlight. These findings contribute to rule out the idea that reduction in trade cost has no impact on regional disparities.

Other interpretations can be made from these results, especially by comparing which shock is more prominent and at which transport cost. Initially, it is not possible to have a definitive answer in comparing each shock strength, since 2SLS and IVRE do not show any difference between parameters estimated. However, expansion shock presents a positive e significant result in the FE by analysing distance to ports, while contraction shock is not significant. On the other hand, distance to internal market might be affected equally after both shocks, since it is not possible to conclude which was stronger. Last, but not least, transport cost to external market seems to present more robust results compared to internal market. Therefore, some conclusions might emerge from this result. First, expansion shock might impact more effectively than contraction shock, but only transport cost to external market. Second, transport cost to domestic market might be affected by any shock. Indeed, similar results are found when not instrumenting human capital. Although there are some suggestions on shock strength and on which market is more affected, it is not possible to have a single conclusion.

One might argue that transport cost may not explain perfectly this demand linkage, since Sao Paulo or any Port regions do not fully shape Brazilian geographical economy.

There are many other regional economic centers and/or regions close to ports, for example, which could have been affected by these demand linkages. As noticed previously, it is possible to investigate it further by using market potential measures, not only internal, but also external. Although it solves this shortcoming, this new approach shows others. These market potential measures are endogenous and instruments are required. As stated previously, population size and distances jointly with government support are used for this purpose, since they may affect regional wages only through market potential measures. Therefore, Equation 3.5 is estimated to fill up the gap which transport cost is not able to do so.[79]

Table 3.4 presents these results, where panel A is responsible to show parameters outcomes, panel B some first stage information and panel C over-identification test.[80] In the former panel, the first two columns present OLS results, where the first with no trade shock and the second using government expenditures. The next three columns show results using two-stage least squares, where the first with no trade shock considered, the second using government expenditures and third using state dummies. Random effects outcomes are presented from (vi) to (viii), following the same structure from 2SLS, which means (vi) no trade effect, (vii) government expenditure and (viii) state dummies.[81] The last two columns show fixed effects results using government expenditure only, with and without trade shock. [82] In panel B, rows are organized as follows. The last four rows report R-squared of each endogenous variable, while the others mention which instrument is utilized in each column. Panel C present over-identification tests of each method used.

[79] Just as a reminder, now the expected sign is the opposite, since bigger markets present higher wages.
[80] As for transport cost results, this table presents only some outcomes for a matter of simplicity. Further results using different controls are shown in Appendix 3.III. In all columns, market potential and human capital measures are instrumented by the earlier mentioned variables, which are distance, population, government support and illiteracy.
[81] Not only human capital measures are divided by Sao Paulo figures in this specification, but also both market potential measures (internal and external).
[82] Hausman test is also performed with outcomes from column v and viii. The hypothesis is also rejected. Therefore, fixed effects present more robust results.

Table 3.4: Panel Data Approach to Equation 3.5

Panel A: Regression Results										
Dependent Variable	OLS		2SLS			IV Random Effects			IV Fixed Effects	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Int. Mkt Pot.	-0.03	-0.02	-0.75	0,11	0.06	-1.23	0.23	0.17	-1.55	-0.59
	(0.018)*	(0.05)	(0.175)***	(0.043)**	(0.027)**	(2.40)	(0.124)*	(0.083)**	(1.25)	(0.60)
Int. Mkt Pot. Cont. Shock		-0.02		-0.12	-0.06		-0.29	-0.24		-0.18
		(0.06)		(0.047)**	(0.031)**		(0.19)	(0.12)**		(0.09)*
Int. Mkt Pot. Exp. Shock		-0.01		-0.12	-0.06		-0.24	-0,19		-0,37
		(0.06)		(0.048)**	(0.031)**		(0.14)	(0.092)**		(0.089)***
Ext. Mkt Pot.	0.25	0.21	1.02	-0.12	-0.06	1.74	-0.26	-0,195	-0.13	-0,14
	(0.015)***	(0.045)***	(0.154)***	(0.046)***	(0.028)**	(1.55)	(0.136)*	(0,09)**	(0.153)	(0.15)
Ext. Mkt Pot. Cont. Shock		0.03		0.14	0.08		0.37	0.30		0.16
		(0.05)		(0.052)***	(0.032)**		(0.209)*	(0.137)**		(0.09)*
Ext. Mkt Pot.l Exp. Shock		0.04		0.13	0.07		0.29	0.24		0.35
		(0.05)		(0.05)***	(0.032)**		(0.157)*	(0.104)**		(0.085)***
Human Capital	0.10	0.10	0.02	0.12	0.09	-0.24	0.09	0.06	0.23	-0.27
	(0.006)***	(0.006)***	(0.06)	(0.038)***	(0.026)***	(0.34)	(0.12)	(0.05)***	(0.19)	(0.11)**
Subsidies	-0.06	-0.06	-0,04	-0.02		0.02	0,06		-0.03	-0.01
	(0.006)***	(0.006)***	(0.008)***	(0.03)		(0.03)	(0.04)		(0.012)**	(0.008)*
Manuf. VA over GDP	0.14	0.14	0.13	0.12	0.14	0.10	0.11	011	0.09	0.10
	(0.006)***	(0.006)***	(0.009)***	(0.017)***	(0.012)***	(0.012)***	(0.017)***	(0.012)***	(0.006)***	(0.007)***
Constant	-0.04	-0.21	0.23	-0.11	-0.07	1.04	1.00	1.15	-11.94	-7.67
	(0.10)	(0.22)	(0.248)***	(0.062)*	(0.033)**	(0.209)***	(0.242)***	(0.499)**	7.97	(4.18)*
Period Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Effect	No	No	No	No	Yes	No	No	Yes	No	No
Observations	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145
R-squared	0.47	0.47	0.20	0.45	0.46	0.13	0.20	0.32	0.23	0.19
Panel B: First Stage Results										
Instruments										
Illiteracy	-	-	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank's Disburs.	-	-	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Share	-	-	No	Yes	Yes	Yes	Yes	Yes	No	No
Distance	-	-	No	Yes	Yes	No	Yes	Yes	Yes	Yes
R-sq HC	-	-	0.14	0.21	0.35	0.25	0.21	0.21	0.26	0.23
R-sq Int. MP	-	-	0.32	0.77	0.54	0.26	0.79	0.79	0.36	0.40
R-sq Ext. MP	-	-	0.35	0.89	0.82	0.49	0.88	0.88	0.37	0.32
Panel C: Results from over-identification Test										
p-value (Chi-sq)	-	-	0.87	0.06	0.88	0.24	0.80	0.90	0.55	0.12

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Before interpreting results, it is important to analyze first stage results of all IV estimation in this section. First, R-square of first stage regressions of all tries are not less than 0.14, where some present really high values, such as 0.89. On one hand, it indicates that instruments are sufficiently correlated to the endogenous variables, but on the other it

raises the question to what extent those instruments may meet the exclusion condition. As it regards to orthogonality condition, over-identification tests presents p-values which encourage the results. Most outcomes presented for over-identification tests are not able to reject the null hypothesis by using 10% level of significance. Those over-identification results are therefore able to provide some evidence that variables used might be valid instruments. However, it is relevant to bear in mind that those statistics are only supportive, and not a definitive answer to whether those are perfect instruments. Summing up, IV results presented in this section are able to satisfy the two most important conditions which are: orthogonal to the dependent variable; sufficiently correlated to endogenous variables.[83]

First, controls continue to behave as previously, by human capital and share of manufacturing sector being mostly significantly positive and subsidies negative. Regarding internal demand, there are different results. At OLS, it seems that internal demand is not able to explain regional disparities in wages. When using IV approach, either through 2SLS and RE, internal market seems to explain those geographical differences. However, results are not corroborated when using FE approach as outcomes are not statistically significant. On the other hand, external demand presents different outcomes. External market potential seems to be positively correlated to regional wages by not using instruments. When they take place, this correlation becomes negative, yet this is not corroborated by estimating with FE. All in all, results seem to corroborate those results using distance as the explanatory variable, but surprisingly not in the most sophisticated model: FE. One possible explanation is that perhaps Sao Paulo is the main driving force to regional disparities in wages, as result using distance (in this case do Sao Paulo) is statistically

[83] Human capital is instrumented by a measure (illiteracy) which raises the question to whether which one is more or less endogenous. One way on dealing with it is estimating 3.5 not using human capital as a control. Omitted variable creates biased results and instruments are needed to tackle this issue. However, results do not present substantive difference, neither at first stage nor at interpreting trade effects, as shown in Appendix 3.I.

significant using FE while outcome utilizing internal market potential (in this case weighting all economic centers) is not.[84]

Trade shocks present even more consistent results compared to Table 3.3. Outcomes presented in Table 3.4 suggest an increase importance of external market while internal market becomes less relevant, including in FE approach. It is safe to conclude that any shock, contraction or expansion, affects regional wages. Agglomeration forces in domestic markets lessened after a reduction in trade cost, while these in regions close to foreign market have their importance strengthened.

In terms of which shock has a greater impact, it seems the difference does not seem to be statistically significant considering the standard deviation in the 2SLS and IVRE approaches. FE approach report results suggesting that expansion shock might impact regional disparities more efficiently, but it is not possible to reach this conclusion by considering 95% confidence interval. It is neither feasible to have a definitive answer on which market is more affected by any of these shocks mentioned, since there are no substantial changes between internal or external market after both shocks.

3.7. Conclusion

This chapter uses Brazilian regional data to test the hypothesis that regional manufacturing wages can be explained by demand linkages and, additionally, how trade shocks can affect these disparities. According to the literature, the most common dispersion force for firms' location decision is competition. Brazil experienced some shocks which have raised competition effect through trade at different periods by contracting domestic activity first then expanding it. This chapter contributes therefore to the literature on

[84]Estimating not instrumenting any endogenous variable show similar results, as shown in Appendix 3.I.

evaluating whether those shocks are able to impact regional wage disparities as well as trying to measure the strength of each shocks and which market might be more affected.

First, regional manufacturing wages are not homogeneous throughout Brazil. They seem to be higher closer to the markets corroborating the hypothesis of agglomeration present in NEG models. Results suggest internal market show domestic market shape regional wages disparities, especially the industry center. External market has the reverse outcome: regions far from external market tend to have higher wages. However, results are not as robust as for internal market.

Economic shocks have moreover changed the importance of internal and external market in explaining regional manufacturing wages. The overall findings show that internal market has turned out to be less relevant, while external market has become more important after both shocks. Comparing which one is more effective and which market might be more affected; outcomes are not conclusive. In spatial terms, Brazilian southern regions appear to have closed the gap to the industry center more after the contraction shock, while northern regions after the exchange rate devaluation. Moreover, coastal regions seem to have benefited the most from both shocks, while regions located in the countryside have fallen behind.

Chapter 4

The Effects of BNDES Loans on Brazilian Manufacturing Firms' Productivity

In previous chapters, some issues related to individuals and firms location decision are investigated. Those decisions might be affected by government policies. One common issue among firms and individuals when they decide where to locate is which region is growing faster. Growth is led by productivity increase. Government interventions to improve firms' productivity may therefore affect regions' growth; therefore more people and firms tend to move to those regions. This chapter investigates whether financial support from a development bank might help firms to improve their productivity.

There is extensive literature examining how firms' performance can be affected by trade policies, especially focusing on the gains from trade. However, there is a shortage of papers addressing whether other government policies, including those designed to develop them, can affect firms' productivity, as evidenced by Grilliches, Klette and Moen (2000) and Criscuolo, Martin, Overman and Reenen (2007). This is not due to a shortage of methods, since other areas have already developed different ways to deal with it. One example is the literature of labour economics which has made a relevant contribution on how to evaluate to what extent government polices affect individuals on their achievements. A survey of this literature can be found in Heckman, LaLonde and Smith (1999). Further work should thus be pursued in evaluating government policies on firms' performance in order to provide empirical evidence for public policy design. This paper

contributes to the literature on firms' performance and public policy by trying to understand to what extent government interventions affect firms' productivity.

Banerjee and Duflo (2005) provide evidence that firms in many developing countries face credit constraints, using a sample of countries which includes Brazil. Therefore, Brazilian firms might be credit constrained, especially for long-term projects. Terra (2003) provides evidence that Brazilian firms are definitely credit constrained. These financial restrictions for long-term projects are considered among the most important market failures in the Brazilian economy as they hamper entrepreneurial efforts of local firms. That is why the Brazilian government provides long-term loans through the Brazilian Development Bank (*Banco Nacional de Desenvolvimento Economico e Social* – henceforth BNDES).[85] The main statutory goal of this institution is to improve Brazilian economic competitiveness without neglecting broader social aspects. BNDES invests in several areas including research and development, infrastructure, export support, and regional and urban development. In the case of manufacturing, BNDES finances long-term projects aimed at the creation of new plants, the enlargement of existing ones, the restructuring and the modernization of production processes, innovation and technological development, and export promotion. Overall, the importance of BNDES in the Brazilian economy is quite sizeable: in 2005 its disbursements reached the value of US$ 20.1 billion, representing 12.3% of aggregate investment.[86]

Although BNDES project analysis involves several dimensions including social and environmental aspects, it is nonetheless interesting to assess their overall impact on the competitiveness of Brazilian firms. The aim of the present paper is to contribute to this assessment by investigating the impact of BNDES activities on the productivity of Brazilian manufacturing firms. Though focused, this is of course a central issue from a

[85] This financial institution presents similar characteristics to the World Bank and the Inter-American Development Bank, although it is uniquely sponsored by the Brazilian Government.
[86] Source: IPEA and BNDES (www.ipeadata.gov.br and www.bndes.gov.br).

policy perspective: “Productivity isn’t everything, but in the long run it is almost everything. ..... Compared with the problem of slow productivity growth, all our other long-term economic concerns - foreign competition, the industrial base, lagging technology, deteriorating infrastructure and so on - are minor issues” Krugman (1992), pages 13 and 18.

To my knowledge, there was no previous study targeting the relationship between BNDES and the productivity of Brazilian firms. Some evaluation of BNDES activities is found in some publications. For its 50th anniversary, a book, BNDES (2002), evaluates many BNDES policies among different sectors. The PhD thesis of Monteiro Filha (1994) investigates how BNDES has supported the consolidation of some sectors within the Brazilian manufacturing industry from 1952 to 1989. Sousa (2003) evaluates the efficiency of BNDES regional polices to achieve its objective of reducing geographical inequalities within Brazil. Capanema (2006) investigates the specific program developed by BNDES targeting the improved competitiveness of the Chemical sector. Reiff, Rocha and Santos (2007) evaluate whether BNDES loans have increased formal jobs in areas supported by the Bank. Pereira (2007) investigates whether supported firms increase employment or not. Torres Filho and Puga (2006) evaluate whether employment and wages increase after a firm has been granted a loan. So far, no evaluation has being done regarding firms’ productivity while and/or after they have been supported by BNDES loans. Amadeo and Gonzaga (1996), Feijo and Carvalho (1994), Ferreira and Rossi (2003), Hay (2001), Muendler (2004), Nassif (2005), Carvalho, Saboia and Salm (1997) as well as Schor (2004) evaluate the productivity of the Brazilian economy after the liberalization process in the 1990s, but they do not investigate the relationship between BNDES and firms’ productivity. Closer to the spirit of the present paper, Lemos, Negri and Negri (2006) evaluate the impact of two funds administered by the Research and Projects Financing Agency (*Fundo de Financiamento de Estudos de Projetos e Programas* – [FINEP]) on

Brazilian firms' productivity and technological upgrading, but they do not consider BNDES loans.

Accordingly, firms are in a situation in which they would like to make investments but credit constraints hamper the implementation of projects. BNDES offers loans that could be used to remove those credit constraints. A key question then becomes: *Are BNDES loans indeed instrumental in relaxing the credit constraints faced by Brazilian firms?* A positive answer requires checking that two things are true. First, granted firms must change their performance after receiving a loan, which can be done by comparing whether these firms changed their productivity after the loan. Second, after receiving a loan, granted firms must perform differently than otherwise identical non-granted firms, which can be done by evaluating both performances. This paper answers these questions by investigating whether firm productivity was affected by BNDES loans.

The empirical results show that on average granted firms do perform better than non-granted firms after being awarded a BNDES loan. However, it is not clear whether this improvement in productivity is associated to BNDES loans. In the most naïve models, a positive correlation is found, but it vanishes when firm heterogeneity is considered (fixed effects). Nevertheless, some interesting results appear when fixed effects are taken into account. Outcomes suggest a positive association for large projects in rich areas and the opposite for small projects in poor areas depending on how many covariates are included. Last, but not least, no effect of BNDES loans on firms' performance is detected by using difference-in-differences approach after comparing with similar firms. Therefore, granted firms perform identically to other similar firms. This may represent that the criterion decision on selecting projects does not focus entirely on economic issues. Other issues, such as social, regional and environmental impacts, might even be as important as

economic aspects (or maybe more) to select the project, due to the amenities that each project creates.

The paper is structured as follows. Section 4.1 begins by describing the financial support offered by BNDES to manufacturing firms, followed by some descriptive statistics in Section 4.2. The methodology and empirical strategy are discussed in Section 4.3. Section 4.4 details the data used in this work. The empirical findings are the focus of Section 4.5. Finally, Section 4.6 offers some concluding comments.

4.1 Overview of BNDES schemes

BNDES provides a wide range of financial tools to support Brazilian manufacturing firms: FINEM, Automatic BNDES, FINAME, Leasing FINAME, International Competition FINAME (BNDES-Exim) and Subscription of Securities.

FINEM ("Financing and Endeavours") is a direct support scheme for projects with financial needs over R$10 million (equivalent to US$ 4.7 million). Projects with financial needs below this threshold are instead supported indirectly through retailing banks under the Automatic BNDES scheme. Both schemes contemplate several categories of expenses covering the creation of new plants, the enlargement of existing ones, the restructuring and the modernization of processes, innovation, and technological development.[87]

Through the FINAME ("Machines and Equipment") and the Leasing FINAME schemes, BNDES supports the acquisition of new domestically produced machines and equipments either buying them (FINAME) or leasing them (Leasing FINAME). Finally, the aim of BNDES-Exim is to provide financial support for exports while the aim of Subscription of Securities is to facilitate changes in firm ownership.

[87] A complete list is available at http://www.bndes.gov.br/english/items_support.asp.

Our focus is on FINEM and Automatic BNDES as they are more focused on supporting the discovery and the implementation of promising projects. Differently, FINAME and Leasing FINAME do not contemplate investments in innovation and technological development. Nonetheless, it is necessary to account for them in order to isolate the role of FINEM and Automatic BNDES. BNDES-Exim and Subscription of Securities have, instead, rather different objectives. FINEM and Automatic BNDES are therefore loans which might affect firms' productivity since their expansion may be guided by improvements in production and/or creation of new and/or more sophisticated goods.

In order to receive any of these two loans, either FINEM or Automatic BNDES, firms need to send a supporting application form with some brief information of their projects to a retailing bank or BNDES itself. The banks evaluate whether their projects are in line with the purpose of the loans mentioned. After having their application approved, firms ought to send a complete and detailed project plan to be evaluated by the financial institution. This project plan is analyzed by investigating whether they are economically viable, what collateral can be used to guarantee the loan and so forth.

The analysis culminates in a formal contract proposal where the terms and conditions of the loan are established, including amount, period, and interest rate. After the negotiation has finished, the loan contract is signed. It is important to note two crucial points here. First, firms do not receive their loan in only one instalment after signing the contract; they receive it gradually according to the development of the project. It is hence during negotiations that disbursements are scheduled over the years of the project implementation. Second, there is a limit for BNDES participation in any project, which is 80% for these two mentioned loans. A project is not therefore fully financed by BNDES, only a part of it is.

Once the loan has been approved, firms receive their first disbursement and those remaining are made after an evaluation of the project's progress. Before the second

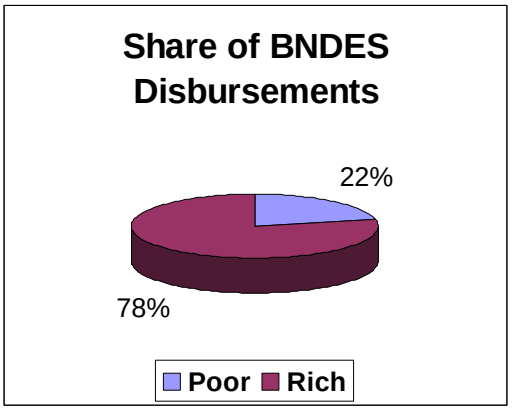
disbursement, the company should prove whether the money of the first disbursement was invested in the items informed by their project plan. Any violation of the loan terms leads to a further investigation and disbursements are interrupted until further explanations are given. If no problems emerge, disbursements continue until the end of the project. Since these are long-term projects, the period between contract signing and the end of disbursements takes on average 4.5 years. It is generally only after all disbursements have been made that firms start to amortization their loans, including interest and principal.

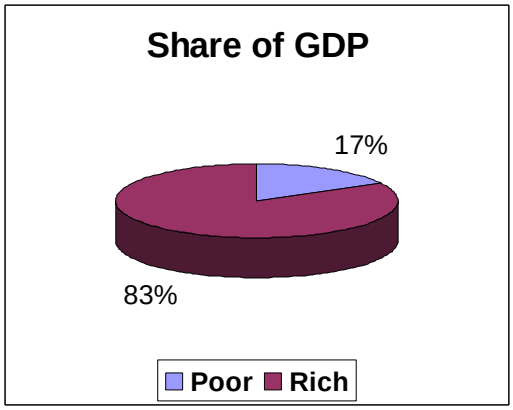
4.2 Descriptive Statistics

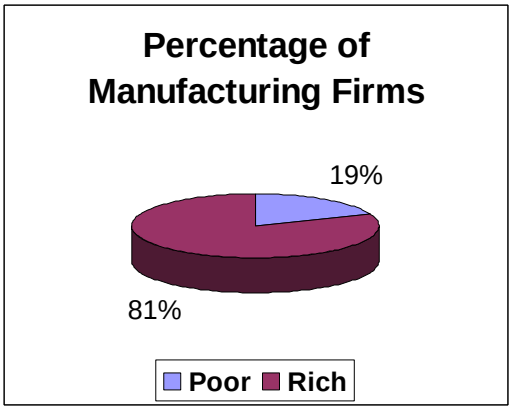
Two channels are explored in order to evaluate how BNDES loans are distributed in the Brazilian economy: rich regions versus poor; manufacturing sector. Information available comprises the last 10 years, which encompasses the years between 1998 and 2007. Since the loans investigated in this work occurred from 1995 until 2003, then only the average from 1998 until 2003 is presented for analysis. Figure 4.1 presents average percentage of disbursement between rich and poor regions in Brazil.

As shown, 78% of BNDES disbursement ended up in rich regions from 1998 until 2003; however proportionately, these areas are much more representative since they represented 83% of the Brazilian GDP and 81% of manufacturing firms during the same period. Therefore, it is safe to conclude that BNDES loans proportionately favour poor regions since they receive a greater share than their actual participation in the economy.

Figure 4.1: Share of Disbursements, GDP and Manufacturing firms between Rich and Poor Regions







Source: BNDES and IPEA

Regarding the manufacturing sector's share, Table 4.1 presents the average share of each sector in BNDES disbursement as well as their share in manufacturing GDP. Generally, BNDES loans reflect sector shares, but two sectors appear to escape from this rule: Transport Equipment; Petroleum Refining and Related Industries. As it regards the latter, although there are many multinationals operating in this sector, the Brazilian government oil company (Petrobras) plays a substantial role.

Table 4.1: Manufacturing Share in BNDES Disbursements and its GDP

Manufacturing Sector	BNDES	GDP
Food And Kindred Products	15%	17%
Tobacco Products	0%	1%
Textile Mill Products	3%	3%
Apparel And Other Finished Products Made From Fabrics And Similar Materials	1%	2%
Lumber And Wood Products, Except Furniture	2%	1%
Paper And Allied Products	6%	4%
Printing, Publishing, And Allied Industries	0%	4%
Petroleum Refining And Related Industries	1%	12%
Chemicals And Allied Products	7%	12%
Rubber And Miscellaneous Plastics Products	2%	4%
Leather And Leather Products	1%	2%
Stone, Clay, Glass, And Concrete Products	2%	4%
Primary Metal Industries	9%	7%
Fabricated Metal Products, Except Machinery And Transportation Equipment	2%	3%
Industrial And Commercial Machinery And Computer Equipment	7%	7%
Electronic And Other Electrical Equipment And Components, Except Computer Equipment	2%	5%
Transportation Equipment	41%	9%
Miscellaneous Manufacturing Industries	1%	3%
Total	100%	100%

Source: BNDES and IBGE

The majority of the investments hence are implemented by this oil company without BNDES support since they have similar access to financial support from international banks as BNDES does. Regarding the former, this sector heavily invested in the development of cars which could run on either ethanol or gasoline (or even a mixture of the both) during this period. Due to the environmental amenities generated by those projects, BNDES provided a substantial part of the financial needs of the majority of these projects. Flex fuel cars are now available for Brazilian consumers contributing to a cleaner environment.

Summing up, results presented here are biased toward poor regions and the transportation equipment sector. On the other hand, rich regions and the petroleum refining and related industries are underrepresented.

The key question is whether BNDES loans are instrumental in relaxing the financial constraints faced by Brazilian firms. A positive answer to this question requires checking first whether granted firms improve their performance after receiving a loan.

Value added per worker (labor productivity) is used as the measure of firm performance.[88] Table 4.2 reports the average values of this variable for granted and non-granted firms in the previous years in which the former received their loans and in the last year of observation, which is 2003. The first three columns present figures of non-granted firms, while information of granted ones is available in the last six columns. For better understanding, a measure of "productivity premium" of granted firms is calculated as the ratio of the average value added per worker of these firms to that of non-granted firms.

Table 4.2: Comparing Productivity of Granted and Non-Granted Firms

Non granted Firms			Only Granted Firms						
Year	Average Productivity	Number of Firms	Year Granted	Number of Firms	Av. Productivity Before	Productivity Premium Before	Productivity in 2003	Productivity Premium in 2003	Annual Increase
1996	51,357	21,533	1997	1,375	73,071	1.42	63,450	2.19	6%
1997	57,440	20,815	1998	1,127	92,699	1.61	82,316	2.84	10%
1998	50,123	22,510	1999	706	118,823	2.37	84,719	2.93	4%
1999	53,034	23,143	2000	801	108,462	2.05	84,452	2.92	9%
2000	43,829	23,284	2001	808	100,307	2.29	87,469	3.02	10%
2003	28,949	23,159	**Average**	**963**	**98,672**	**1.95**	**80,481**	**2.78**	**8%**

Source: Brazilian Statistical Institute and BNDES

The table shows that granted firms are generally more productive than non-granted. One year before loans were approved, productivity premium of granted firms ranges between 1.42 for firms granted in 1997 and 2.37 for those in 1999. On average, firms selected to be financed are twice as productive as those not. Moreover, outcomes show that BNDES loans select even more productive firms over the years, since the premium increases from 1.42 to 2.29.[89] The last column shows that the premium grows through time from the years loans were awarded in all cases. On average, productivity premium of granted firms has increased by 8% annually over non-granted.

[88] Potential shortcomings of using this measure are discussed in Section 4.4.
[89] The only exception is in 1999, when productivity premium achieved its maximum.

4.3 Empirical Strategy

Results from descriptive statistics seem therefore consistent with the idea that BNDES financial support has helped Brazilian firms to relax their credit constraints since granted firms performed annually 8% better than other non-granted firms. This is however very rough evidence that may not survive closer econometric scrutiny. In order to be as accurate as possible, five methods are implemented, which are presented in increasing order of sophistication.

4.3.a Method 1 – Structural break

The starting point is to check whether the productivity improvement of granted firms is associated with BNDES loans after controlling for other observable characteristics of firms. For that, the simplest specification is the following:

$$y_{it} = X_{it}\beta + \delta D_{it} + \varepsilon_{it} \qquad (4.1)$$

where y_{it} is value added per worker, D_{it} is a dummy variable for all years after the loan being given, X_{it} is a vector of control variables, ε_{it} is the error term and the rest are parameters to be estimated. In particular, the impact of BNDES loans on firm productivity is measured by the estimated value of δ. The control variables include various firm characteristics: age, wage, number of employees, their skill and education levels, market share, revenues, financial status, exports, intermediate imports, capital imports, multinational status and FINAME loans.[90]

One shortcoming of using value added per worker is that it also captures improvements in productivity which is created by investments in capital. However, some controls included are able to capture these things, such as capital imports and FINAME

[90] A full description of these variables is available in Appendix 4.I.

loans. The latter represent most of the improvements by domestic capital goods, since more than 80% of all manufacturing firms have financed their capital investment by borrowing from FINAME during the period analyzed. While the former represent all capital goods imported by any Brazilian firm at each year encompassed in this period. In other words, it is a record of how much capital goods were imported by any firm in every single year during the period analyzed. Therefore, investments in capital goods are capturing this part of increase in productivity.

There are some potential problems with specification (1) that may lead to biased estimation of the coefficients. First, there may be unobserved characteristics affecting both dependent and independent variables. Second, the linear relation between X_{it} and y_{it} may not be the most appropriate. Third, X_{it} and D_{it} may be endogenous.

4.3.b Method 2 – Grant identifier

One way to deal with the first potential bias of Method 1 is to introduce another dummy variable, which takes value 1 if a firm has ever been supported by BNDES loan during the period of observation, and 0 otherwise. The specification then becomes:

$$y_{it} = X_{it}\beta + \delta D_{it} + \alpha EVER_i + \varepsilon_{it} \qquad (4.2)$$

where *EVER* is the new dummy variable that captures unobserved time invariant differences between granted and non-granted firms. In other words, if there were some unobserved characteristics influencing the performance of granted firms both before and after receiving a loan, the new dummy would pick them up. However, if there were additional unobserved heterogeneity across firms and such heterogeneity were unrelated to their granted status, specification (4.2) would still lead to biased estimation.[91]

[91] Firm heterogeneity per se does not create bias. However, as it is correlated to the dependent variable, it does lead to biased estimation.

4.3.c Method 3 – Fixed effects

One way of dealing with unobserved heterogeneity across firms unrelated to their granted status is to use a fixed effect specification. This can be derived from (4.1) by writing the error term as:

$$\varepsilon_{it} = \gamma_i + \lambda_t + \nu_{it} \qquad (4.3)$$

where γ_i is the fixed effect capturing all unobserved firm characteristics that are time-invariant and λ_t for any unobserved time-variant characteristic.

4.3.d Method 4 – Productivity Growth

In order to evaluate whether BNDES loans affect productivity growth, it is useful to go from a specification in levels to a specification in differences:

$$\Delta y_{it} = \Delta X_{it}\beta + \delta D_{it} + \Delta\varepsilon_{it} \qquad (4.4)$$

which focuses on the relation between BNDES loans and the growth of firm productivity. It is important to notice that dummy variables are still in levels yet productivity in changes.[92]

Although these four methods may solve the problem of biased estimators, all still assume a linear relation between control variables and firm performance. Moreover, they may reveal an association between BNDES loans and firm performance but not an effect. A way to relax the linearity assumption by finding a counter factual group of non treated firms and to estimate the BNDES effect is through matching and difference-in-differences approach estimation, respectively.

[92] One might argue that estimating FE specification by first differencing might be useful. However, results on this method presented in Section 4.5 show no association. Then, improving the method might not reach different conclusions.

4.3.e Method 5 – Propensity score matching

In order to assess the impact of BNDES schemes on firm performance, ideally one would like to compare granted and non-granted firms which were identical in any respect when loans were awarded to the former. Propensity score matching (PSM) allows one to pursue such an ideal comparison by matching each granted firm with a non-granted exhibiting similar observable characteristics. This artificially generates a "control group" of non-granted firms to be compared with the "treated group" of granted firms in order to assess the impact of "treatment" by BNDES schemes. The comparison can then be made in terms of productivity levels.

There are different ways to implement PSM[93] and this paper uses the Kernel method.[94] This method creates a counter factual group by pairing each granted firm with all non-granted firms weighted by how similar they are. In other words, weights are allocated to non-treated firms according to how alike they are compared to supported ones. The main advantage of this method is that all information of the sample can be used, which will lead to a non control group of more than 15 thousand firms.

Matching is based on the following pre-treatment characteristics, such as productivity, age, number of employees, average salary of employees, market share, total revenues, percentage of highly skilled workers (undergraduate and above), sector, location, educational level of employees (years of schooling), ratio of financial costs over total revenues, ratio of exports over total revenue, ratio of imports of capital goods over investment, ratio of imports of intermediaries over cost of production, and multinational status.[95]

[93] PSM is used by Negri, Lemos et al (2006) to evaluate the impact of FINEP on firm productivity and R&D investment. See also Arnold and Javornik (2005) for a detailed implementation of PSM on foreign investment in Indonesia.
[94] Further information of this method is presented in Appendix 4.III.
[95] More details of each variable is available in Appendix 4.I

The evaluation of the treatment effect is performed by differences-in-differences according to the following specification used in Bronzini and Blasio (2006):

$$y_{it} = \beta BNDES_i + \sum_t \alpha_t D_t + \sum_t \delta_t (BNDES_i \cdot POST_t) + X_{it}\gamma + \varepsilon_{it} \quad (4.5)$$

where y_{it} is again value added per employee, $BNDES_i$ is a dummy variable indicating whether the firm received any loan in the period of observation, D_t is a year dummy, $POST_t$ is a set of dummies for each year after receiving the loan, and X_{it} is the vector of control variables. The parameter of interest is delta which estimated value measures the impact of BNDES schemes on firm productivity over time. Note that the estimation of (4.5) allows one to assess not only whether BNDES loans affect firm productivity but also when the impact materializes.[96]

Before data description, three additional comments are in order. First, to evaluate the impact of FINEM and Automatic BNDES loans, it is necessary to isolate it from the effect of FINAME. This cannot be achieved by eliminating all firms that received FINAME as most firms in the sample received different types of loans jointly. For instance, during the period of observation, 26 per cent of firms received FINEM or Automatic BNDES loans while more than 80 per cent received FINAME support.

Since the information available is only whether the firm got FINAME or not during the whole period (and not when), there is only one way to control for this effect. A dummy variable for firms receiving FINAME loans during this period is created. Since it is a time-invariant characteristic, it cannot be used in Methods 3 and 4.[97]

[96] It is also feasible to evaluate time in previous specifications, but they would say only when association occurs. In this method, knowing when treatment effect materializes seems to be more interesting.
[97] If at least information on when the firm has received this loan was available, then another alternative could have been creating a dummy after they received the loan. Further information of this loan may improve results, but it seems not to matter much as outcomes in Section 4.4 show.

In addition to FINAME, there is a second issue worth discussing. Some firms were granted FINEM or Automatic BNDES loans more than once in the period of observation. Although more than 90% of all granted firms received it only once, repeated treatments may distort the overall picture. To tackle this problem, there are two possible strategies depending on methods. For Methods 1 to 4 ("regression analysis"), the dummy D_{it} takes different values: one after the first treatment, two after the second treatment, three after the third treatment, and so on.[98] Differently, for Method 5 any firm granted FINEM or Automatic BNDES loans more than once is removed from the sample in order to eliminate any double treatment.

4.4 Description of the dataset

Data are drawn from a variety of sources already used by Lemos et al (2006). The dataset combines information from: the Annual Industrial Research (*Pesquisa Industrial Anual* – [PIA]) of the Brazilian Institute of Geography and Statistics (*Instituto Brasileiro de Geografia e Estatística* – [IBGE])[99]; the Annual Social Information Report (*Relação Anual de Informações Sociais* – [RAIS]) of the Ministry of Labor; the Foreign Trade Secretary (*Secretaria de Comércio Exterior* – [SECEX]) of the Ministry of Industrial Development and Foreign Trade; the Foreign Capital Census and the Central Bank Register of Brazilian Capital Abroad of the Brazilian Central Bank; BNDES itself.[100]

[98] One implicit assumption is that every treatment has the same effect. However, one might expect that the first might be more important than any further treatment. Therefore, creating other dummies for different treatments might be more suitable. Nevertheless, results presented in Section 4.4 show that this may not make too much difference.

[99] This is our main data source, since it contains the majority of the variables useful for this analysis, including productivity.

[100] The construction of the dataset has followed procedures that guarantee the confidentiality of information so that individual datum cannot be related to any specific firm.

BNDES data are used to identify firms given loans. Information is available from 1995 to 2003.[101] In this period 539 firms received FINEM loans and 8,505 firms borrowed under the Automatic BNDES scheme. In the latter case, it is not possible to use all firms since some of them are not available from PIA. The reason is that PIA covers all firms with more than 30 employees and some firms granted Automatic BNDES are smaller than that. This reduces the number of firms granted Automatic BNDES in our sample by half.

Our sample is further reduced by three additional issues. First, the focus of this paper is on *manufacturing firms* while BNDES records concern all *manufacturing projects*. In other words, they report manufacturing projects by non-manufacturing firms (e.g., those of large food retailers investing in the development of their own brands) and do not cover non-manufacturing projects of manufacturing firms (e.g., in agriculture). Hence, we end up evaluating only the performance of manufacturing firms granted loans to implement projects in the manufacturing sector. A second issue is related to mergers. For example, if Firm A received a loan in 1997 and in 2000 merged with Firm B creating a new Firm C, and then the initial loan would be registered to firm C. As the past records of Firm C are impossible to reconstruct, all information on loans projects granted to firms like A and B are dropped. Finally, there is a time lag of generally two to three years before a firm enters the census part of PIA. [102] Hence, some granted firms with more than 30 employees are not recorded by PIA at the moment they receive BNDES loans.

In the end, data are left for 240 firms granted FINEM loans and more than 2,000 firms granted Automatic BNDES from 1996 to 2003. Around 15,000 non-granted firms are available to construct the counter factual group. Control variables are reported in Appendix 4.I.

[101] Data on 1995 is used only to exclude any firm which received treatment in that particular year.
[102] IBGE receives information of firms' size (number of employees) from a particular year only at the end of the following year. Thus, any new firm for the census part will provide information only after two or three years having become eligible to be computed.

Although this treated group has a reasonable size, potential drawbacks of conducting a policy analysis with this treated group should be addressed. First, any of these firms (granted or not) might be affected by other government interventions apart from BNDES loans. Second, there may be a time lag for any impact to be detected, since outcomes do not necessarily appear immediately after the loan has been granted. For the former problem, it will be assumed that BNDES loans are the main type of subsidies to affect firms' productivity, since this is the largest financial institution in Brazil offering loans for long-term projects. For the latter shortcoming, BNDES loans encompass six months before the project may be considered eligible for analysis. Then when the loan is approved, the project is already being implemented, and considering the effects on the current and subsequent period is not a strong assumption.[103]

4.5 Empirical Results

In order to see where the policy may have a stronger impact, two channels are explored: location of the project (poor or rich regions); and project size (small or large projects).[104] Therefore, results presented in this paper show not only how these loans affect Brazilian manufacturing firms, but also whether small projects have a higher impact and/or whether firms in poor regions benefited more from these financial supports.

In order to evaluate by the simplest method (Method 1), Table 4.3 presents results of δ from this approach, which is done by just including a BNDES dummy.[105] The first column lists different samples in each row by distinguishing between small and large

[103] Firms are credit constrained to implement the whole project. However, BNDES does not finance the full project and rather a part of it as explained in Section 4.1. Some projects might start before credit has been given.
[104] Rich areas are considered the South and Southeast regions in Brazil, where North, Northeast and Middle-West are named as poor regions. Regarding small and large projects, any projects supported by Automatic BNDES is considered a small project, while any project supported by FINEM is called a large project.
[105] A table containing results of the remaining parameters estimated are presented in Appendix 4.II.

projects as well as whether the firms locate in a poor and/or rich area. The first row of results presents outcomes from all projects supported by the Bank. In the second and third rows, projects are distinguished by size with the second considering only small projects while the third only large projects. Results regarding the location of projects are shown in rows four and five which the latter is related to projects developed in rich areas, while the former in poor areas. The last four rows present results by location and project size. Thus, when it states "Small Projects in Poor Areas," it means only small projects developed in poor regions are considered.

The columns represent how many controls are included in the estimation process. Only controlling for FINAME is presented in the first column, while the second shows results when adding sector and region dummies as controls. The last column exhibits results with all controls mentioned for matching, apart from productivity measure (the dependent variable).[106]

Table 4.3: Results Using Method 1 – Structural Break

Dependent Variable: VA per Employee	Only FINAME	Plus Sector and Region Dummies	with all controls
Full Data	0.47***	0.44***	0.06***
Small Projects	0.41***	0.40***	0.07***
Large Projects	1.08***	0.97***	0.06***
Poor Areas	0.62***	0.51***	0.08***
Rich Areas	0.43***	0.42***	0.07
Small Projects in Poor Areas	0.57***	0.45***	0.08**
Small Projects in Rich Areas	0.38***	0.38***	0.07***
Large Projects in Poor Areas	1.10***	0.94***	0.09
Large Projects in Rich Areas	1.09***	0.98***	0.06

*** means 1% of significance, ** means 5% and * means 10%.

By using the most naïve model (Method 1), results suggest a positive association between receiving BNDES loans and firms' performance, even after controlling for all possible covariates as shown in the first line. Distinguishing by large and small projects, it appears that large projects are more positively correlated with government policy at first,

[106] A list of all controls is listed in the Appendix 4.I.

since large projects have double the effect of small projects using only FINAME dummy, as shown in the first column. However, results suggest no significant difference between small and large projects after controlling for all characteristics as the last column shows. Looking at where this association is stronger, in other words rich versus poor areas, outcomes suggest that BNDES loan dummy is positive and significant in poor regions, but not in rich regions, as the last column states.

Although policy dummy is positive and significant in large projects solely, no evidence is found when this result was further investigated by distinguishing between rich and poor regions, as shown in the last column of the last two rows. On the other hand, there is a positive and significant association between small projects and this financial support regardless of whether it is a poor or rich area.

These results suggest that small projects might have a better performance than large ones. This outcome reveals the small and medium-sized firms might be credit constrained and they are those which could be better benefited by government policies.[107] Regarding large projects, although they seem to be positively associated to BNDES support, when distinguishing between rich and poor areas, results suggest no correlation at all. In other words, firms that receive these loans seem to have the same performance as those which have not received. Nevertheless, these results should be interpreted with care, since parameters are biased due to the fact that there might be unobserved differences between granted and non-granted firms. Controlling for this heterogeneity between them, Method 2 captures this distinction of being granted or not. Table 4.4 presents outcomes for this method using an identical structure of Table 4.3 rows and columns.

[107] No restrictions exist for large firms to implement small projects; however, the majority of these small projects are run by small and medium-sized firms.

Table 4.4: Results Using Method 2 – Grant Identifier

Dependent Variable: VA per Employee	Only FINAME	Plus Sector and Region Dummies	With all controls
All	0.13***	0.13***	0.05***
Small	0.13***	0.13***	0.06***
Large	0.09***	0.11***	0.04
Poor	0.22***	0.19***	0.05
Rich	0.11***	0.12***	0.06***
Small & Poor	0.25***	0.21***	0.06
Small & Rich	0.11***	0.12***	0.06***
Large & Poor	0.03	0.12	0.01
Large & Rich	0.10*	0.13***	0.06

*** means 1% of significance, ** means 5% and * means 10%.

Results do not present substantial differences to the previous method. Overall, firms' productivity is positively associated to BNDES loans even controlling for all covariates, as shown in the last column of the first line. However, this is only evidenced by small projects all over Brazil or in rich areas regardless of a project's size. In other words, after controlling for unobserved characteristics which could be distinguished between granted and non granted firms, such as management aspects, a firm financially supported in a small project by the Bank performs better than another not supported. The same type of conclusion can be drawn by considering results for loans to firms in a rich area, where granted firms tend to develop more than non-granted ones. On the other hand, productivity of firms implementing large projects in Brazil as a whole or in poor areas regardless of projects' size is not positively associated to BNDES loans. This suggests that firms financially supported in large projects do not present better performance after controlling for unobserved characteristics which could be distinguished by being granted or not. A similar interpretation can be made for projects in poor regions, where two firms with the same unobserved characteristics of being granted or not will perform alike even after one is granted a loan and the other not.

Taking into consideration not only the size of the project but also where it is located, results back up that only small projects in rich areas developed by a granted firm are positively associated to the Bank's loans. Therefore, unobserved characteristics

distinguishable by being granted or not ripped off any positive correlation on all other three combinations, which are: poor and small projects; poor and large projects; and finally, rich areas and large projects.

Again, results of Method 2 are biased since not all unobserved characteristics may be represented by being granted or not. Therefore, Method 3 considers all unobserved features between firms by including firms' fixed effects. Table 4.5 uses the same row structure Tables 4.3 and 4.4 rows, but columns differ. Since no time invariant characteristic is able to be used here, then all dummies (FINAME, state and manufacturing sector) used previously are not included.[108] Therefore, first column uses no control at all while the second uses all time variant controls.

Table 4.5: Results Using Method 3 – Fixed Effects

VA per employee	No Controls	All covariates
All	-0.01	-0.01
Small	-0.03***	-0.01
Large	0.09*	-0.00
Poor	-0.04	-0.05***
Rich	-0.01	-0.00
Small & Poor	-0.07***	-0.06
Small & Rich	-0.02	-0.00
Large & Poor	0.06	-0.06
Large & Rich	0.09*	0.01

* means 1% of significance, ** means 5% and *** means 10%.

Outcomes became substantially different after allowing for unobserved firm heterogeneity. First, the general positive correlation between loans and productivity vanishes after considering this more realistic assumption. This suggests that a firm's unobservable characteristic may explain most of the differences and evolution of productivity level. This result supports the fact that granted firms were always more productive before having their loans approved as shown in Table 4.1. Therefore, the selection process of deciding which firms should be granted encompasses all firms' unobservable characteristics, such as management profile. In other words, firms'

[108] The amount of FINAME loans is also used as control here. However, results remain the same. For simplicity, only outcomes with and without all controls are presented in this section.

productivity is not associated with BNDES loans after controlling for these unobserved characteristics.

Without controls, BNDES dummy is positive and significant for large projects, but negative and significant for small projects. This result suggests that firms implementing small projects decrease their productivity after the loan. On the other hand, firms financially supported for large projects do perform better than another not granted. Nevertheless, these results are not backed up when all covariates are taken into account. The single example of a significant relationship after controlling for all covariates is for poor regions, which suggests a negative association between firms' productivity and financial support, even though it is not observed without controls.

When trying to identify where these results are more evident, it is feasible to conclude some facts. First, only large projects in rich areas seem to show a positive relation between productivity and BNDES support only without controls. This means that large projects may be prosperous only in rich areas, where economies of scale may help them to become more productive. Second, the negative association between banks' loans and productivity is only noticed by small projects in poor areas. Therefore, firms implementing small projects in poor areas do perform under the achievements of similar firms. It is nevertheless important to mention that these results (positive and negative) are not supported by taking into consideration all possible covariates. One possible explanation for that might be that BNDES loans may have affected also some control variables such as revenues. Therefore, this might be misleading the results towards no association or vice versa.

In order to evaluate whether the loans is associated with the increase of productivity (not level), Method 5 provides an alternative. Table 4.6 shows results of this method with the same structure of Table 4.5.[109]

Table 4.6: Results Using Method 4 – Productivity Growth

VA per Employee	No Controls	with all covariates
All	-0.00	0,01
Small	-0.00	0,01
Large	0.01	0,00
Poor	-0.01	0,00
Rich	-0.00	0,01
Small & Poor	-0.01	0,00
Small & Rich	-0.00	0,01
Large & Poor	-0.01	0,01
Large & Rich	0.01	0,01

*** means 1% of significance, ** means 5% and * means 10%.

It is possible to see that no association is found in rich or in poor areas and neither between large and small projects. Even distinguishing between location and project size, outcomes do not present any single correlation in this method. Therefore, it is feasible to conclude that even finding some association between levels of productivity and BNDES loans, there is not a single correlation between changes of productivity (increase or decrease) with the financial support.[110]

Although these four methods highlight whether there is any association between the government policy and firms' performance, they cannot assess any effect of BNDES loans on firms' productivity. Another issue resides in the fact that any non linearity between BNDES loans and firms' productivity cannot be detected. Those points can be answered by using PSM and difference-in-differences.

As detailed in Section 4.3, the aim of PSM is to create a counter factual group of non-granted firms ("control group") which are identical in any respect to the granted firms

[109] Just to remind that FINAME, state and sector dummies are not feasible in this method due to the same reasons used in method 3. FINAME amount was also used, but there were not substantial changes. For simplicity only results with and without controls are presented here.

[110] As stated in Section 4.3.d, estimating FE specification by first differencing might have been an option. However, results on methods 3 and 4 suggest that no association may be found as well.

("treated group") when they, the latter, received loans ("treatment"). This controls for potential nonlinearities in the relation between control variables and firm productivity.

The implementation of (4.5) through PSM is not possible for all years. On the one hand, some time should be allowed in order to check the full effect of the treatment. Since Automatic BNDES and FINEM last for at least five years,[111] then it deserves a period beyond a five-year horizon to check its impact. This allows the model to check BNDES' effects not only during but also after the treatment. Given the time spanned by our dataset (1996 to 2003), that is clearly not feasible for loans granted from 1998 onwards. On the other hand, to build the control group for firms treated in a certain year, one needs to have at least one previous year of observation to be used for PSM. Hence, the impact of BNDES schemes can be scrutinized through PSM only for firms granted Automatic BNDES and FINEM loans in 1997, which therefore define our treated group. Moreover, to avoid concerns regarding repeated treatment, all firms treated more than once from 1995 to 2003 were excluded. Lastly, the construction of the control group is based on all firms active in 1996 which were never granted loans during the entire period of observation from 1995 to 2003 and has survived during the whole period. In the end, 291 treated and 15,127 non-treated firms are left to perform PSM.[112]

After matching non-treated firms with treated, difference-in-differences approach of Equation 4.5 is used in order to infer whether any effect of BNDES loans on firms' productivity exists. As in the previous methods, different specifications are tried distinguishing between small and large projects as well as poor and rich regions with or without controls. For parsimony, Table 4.7 reports only the results for the specifications in

[111] The average time of all disbursements from both loans is around 4.5 years.
[112] This PSM was done by estimating firms' probability to get a loan from BNDES by using a probit model with the characteristics mentioned in the Appendix 4.I (all of them from the year before granted year, in this case 1996). After having this probability, weighs are given to non-treated firms by how similar they are to granted ones. More details are presented in Appendix 4.III.

which only FINAME or all controls are introduced. Rich and poor areas are considered jointly as well as separately.[113]

Table 4.7: Results Using Method 5 for Any Project – PSM plus Dif-in-Dif

Dependent Variable	All Regions		Poor Regions		Rich Regions	
Value Added per Worker	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Policy effect in 97	-0.07	-0.08	0.09	0.09	-0.09	-0.07
	(0.12)	(0.10)	(0.27)	(0.14)	(0.13)	(0.11)
Policy effect in 98	0.13	0.12	-0.24	-0.24	0.19	0.16
	(0.13)	(0.11)	(0.44)	(0.33)	(0.13)	(0.11)
Policy effect in 99	0.01	-0.17	0.57	0.37	-0.07	0.03
	(0.11)	(0.09)	(0.44)	(0.31)	(0.11)	(0.12)
Policy effect in 00	0.01	-0.03	-0.24	-0.19	0.04	-0.08
	(0.13)	(0.11)	(0.29)	(0.21)	(0.15)	(0.12)
Policy effect in 01	-0.09	-0.04	0.20	0.23	-0.14	-0.10
	(0.15)	(0.13)	(0.30)	(0.20)	(0.17)	(0.12)
Policy effect in 02	-0.04	-0.02	-0.16	-0.21	-0.02	-0.01
	(0.15)	(0.12)	(0.33)	(0.20)	(0.17)	(0.12)
Policy effect in 03	0.06	0.04	0.00	-0.11	0.08	0.05
	(0.13)	(0.09)	(0.33)	(0.22)	(0.13)	(0.13)
Finame	0.40	0.12	0.48	0.02	0.39	0.09
	(0.04)***	(0.03)***	(0.10)***	(0.07)***	(0.04)***	(0.01)***
Age		0.05		0.00		0.02
		(0.02)**		(0.07)		(0.01)**
Skill		-0.09		0.26		0.40
		(0.20)		(0.50)		(0.08)
Wage		0.35		0.23		0.41
		(0.02)***		(0.05)***		(0.01)
Schooling		0.10		0.20		0.03
		(0.06)		(0.14)		(0.02)
Number of employees		-0.62		-0.63		-0.59
		(0.02)***		(0.03)***		(0.01)***
Market Share		0.44		-11.22		1.42
		(1.35)		(5.34)**		(1.81)
Revenues		0.72		0.78		0.66
		(0.02)***		(0.03)***		(0.01)***
Financial Status		-0.70		-1.44		-0.53
		(0.28)**		(1.24)		(0.05)***
Export		0.11		0.52		0.15
		(0.07)		(0.12)***		(0.04)***
Intermediaries import		0.19		0.79		0.01
		(0.11)*		(0.28)***		(0.07)
Capital import		0.19		0.48		0.11
		(0.05)***		(0.14)***		(0.06)**
Number of Obs.	110,275	109,544	15,068	14,989	95,207	82,776
R-squared	0.02	0.38	0.02	0.47	0.02	0.36

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

[113] Other results, such as including small and medium projects, are presented in the Appendix 4.II.

Regarding controls, it is relevant to mention that most of them present the expected sign. For example, the covariate variable used to capture FINAME is always positive and significant at the 1% level despite where: rich and/or poor areas. Capital import is positive as well. These two results are an indication that these control variables of investment in physical capital is reasonably captured by those measures. Therefore, the shortcoming of using only labor productivity may be in some sense being overcome by using these measures. Moreover, results suggest that those controls appear to capture any improvement in productivity related to capital investments. Other controls present positive and expected signs, such as wages and revenues at all regressions. Another interesting result is that employment is negatively related to productivity as expected, since the more employees, the lower is the labor productivity.

Focusing on the BNDES loans, despite being a poor or rich region, results show that no impact is found by those loans on firms' productivity, since not a single parameter estimated appears different from zero. Therefore, it is not possible to reject the hypothesis that the Bank has any impact, negative or positive, on firms' productivity. It is important to notice that the impact is evaluated not only while the project is being implemented, but also some time after it has finished. It is assumed that all projects end their implementation after five years (the average loans' duration at this bank), which enables this paper to see the impact two years after implementation has been completed. Even considering this ex-post period of projects, no significant result is found.

This outcome is not an isolated case in the literature. Criscuolo et al (2007) investigated the effect on industrial policy in the UK. Their results show no significant impact on firms' productivity, even though they found effects on employment and investment. Considering that Pereira (2007) and Torres Filho and Puga (2006) have also

found that BNDES loans impact positively on employment in granted firms, then finding no impact on labor productivity might be related to those previous findings.

There are two other time issues, one before and another after the loan itself, which might help to understand this result (no impact). Major adjustments to become more productive may have occurred before applying for financial support, since BNDES policies select mainly successful enterprises, as Table 4.2 reports. Therefore, firms might have already become more productive in order to be eligible for this financial support. After getting it, their performance does not change substantially. As a result, BNDES effect on firms' productivity might be prior to the announcement of the loan approval. On the other side, ex-post evaluation of some similar institutions, such as in the European Bank for Reconstruction and Development, occurs only two years after the end of any project, as pointed out in EBRD (2006). In order words, it might be too soon to see any impact on firms' productivity within two years after the project has finished.

4.6 Concluding remarks

Overall, granted firms perform better after receiving BNDES loans. On average, firms improve their productivity by 8% annually after being treated compared to non-treated firms. When examining whether any association between firms' performance and BNDES loans, they are found solely in naïve models. Moving towards more sophisticated models, this association vanishes. In naïve models, it was possible to see some positive association between firms' productivity and BNDES loans.

Additionally, no evidence was found related to effect of BNDES financial support on firms' productivity, as evidenced by other similar papers. Explanations on why not include when the effect really occurs (before or after) and whether an employment increase cancels off any labour productivity effect. Some further conclusions can be reached by

those outcomes. Initially, as firms perform better after receiving a loan, then economic aspects of each project appear to be relevant as one criterion decision for loans approval. As identical non-granted firms perform similarly to those supported, then perhaps the criterion decision on selecting projects does not focus entirely on economic issues. Other issues, such as social, regional and environmental impacts, might even be as important as economic aspects (or perhaps more) to select any project, due to the amenities that each project creates.

This no treatment effect results may be similar to a controversial issue of educational grants, where academics question whether scholarships improve research's quality or higher skills researchers are selected for financial support. Despite that, further work should be pursued to investigate whether these loans can eventually impact firms' productivity or not, especially by using total factor productivity instead of labour productivity.

Chapter 5

Conclusion

Throughout this study, fundamental questions over individuals' and firms' location are addressed, especially focusing on the determinants of each decision based on economic, social, and policy issues.

Chapter 2 evaluates whether amenities or disamenities, especially violence, are correlated with urban-urban migration flows between Brazilian cities. Not only violence was used to explain this, but also other disamenities, such as population density, and other amenities, such as entertainment varieties as well as many social and economic variables.

At first, a restricted gravity equation focusing on destination characteristics is estimated in order to evaluate whether people have changed dwelling towards safer cities. Results confirmed that migration flows are negatively correlated to violence; hence people have moved to live in cities with lower crime rates. Considering that only one type of crime is tested (homicide rates), population size results suggest that other types of violence may be influential for migration decisions. This outcome is backed by the violence results, which demonstrates that migrants have decided to live in less dangerous regions. Population density, which is a proxy for another disamenity such as pollution and congestion, does also present a negative correlation to migration flows. This means that migrants are avoiding not only less populated places, as population size results suggest, but also less densely populated cities. On the other hand, entertainment options, used as a proxy for other amenities, seem to be positively related to migration flows. Overall, results using only destination attributes suggest what was expected, which is: migrants avoid places with disamenities and move towards regions with higher amenities.

One issue further investigated was whether short- versus long-distance migration flows differs. The classic example for the migration literature is between wages and housing costs. For short distance, housing costs might be more important, while wages may be more relevant for long distances. With regards to amenities, one might also think that they may differ depending on the distance. Population density and entertainment options show similar results which are negative for the former and positive for the latter. Results suggest moreover that they matter more in short distance compared to long distance. Violence shows a pattern similar to those. The variable which presents the most interesting outcome is population size. For short distances, it is positively related to migration flows which means that people move towards bigger cities. The contrary occurs with long distances. One possible explanation for this outcome may be that migrants to surrounding cities are also impacted by these disamenities, such as other types of crime, at their origin cities; therefore, moving towards a bigger city does not matter.

Chapter 2 also explores whether origin attributes are correlated to migration flows by using a traditional gravity equation with origin and destination characteristics as explanatory variables. At first, amenities and disamenites only at destination seem to be correlated with migration flows since origin attributes do not present robust results. Population size presents a positive association between origin measures and migration flows, while destination measures show a negative association. This outcome is an indication that criminality at origin may be repelling people towards other places, where lower levels of violence are observed. The violence measure (homicide rates) reinforces this argument since they also present the same sign, which is negative at destination and positive at origin. Origin attributes show, however, much more robust results than destination.

Violence is a social amenity which might have neighbouring effects due to the social interaction among criminals. Lastly, Chapter 2 checks whether the spatial lag of violence is associated to migration flows. Two spatial matrices are used separately: contiguity, and nearest neighbours. Regardless which matrix is used, the restricted gravity equation using only destination attributes shows that violence at local and surrounding areas is negatively related to migration flows. Migrants thus try to avoid dangerous regions when they decide where to move. By using the traditional gravity equation, origin characteristics appear to be positively related to migration flows and destination turned out weaker. Therefore, origin attributes seem to be more relevant for migration decisions than destination decisions.

Overall, Brazilian migration flows seem to be associated with social amenities and disamenities, especially violence, as outcomes suggest. These results provide a contribution to the literature on migration since they provide further evidence that not only economic reasons drive people to migrate but also social amenities and disamenities. Further work using individual data should be pursued in order to investigate whether violence was actually a proper reason to migrate.

Aside from individuals' movements, firms also decide where to locate. Their decision on where to establish a production plant takes into consideration two marginal costs: transport costs, and wages. According to the literature, these two costs are negatively related. Chapter 3 investigates whether wages are a decreasing function of transport cost, as well as whether they are positively associated with market potential. Due to the two trade shocks which Brazil experienced during the 1990s, it was also possible to evaluate whether these shocks were able to affect regional wage inequalities, as the literature suggests. The particularity of these shocks, moreover, provides further analysis on which dispersion force is more important to impact these regional disparities.

Using Brazilian regional data, Chapter 3 examined initially which regions geographically have been affected by these shocks. The map suggests southern regions have closed the gap with the industry center after the contraction shock, a unilateral liberalization process. On the other hand, northern regions seem to have improved in relation to Sao Paulo after the real exchange rate devaluation at the end of the 1990s (referred to as expansion shock). These geographic interpretations suggest that those trade shocks might have affected regional disparities in Brazil.

An econometric specification is tested by trying to probe whether wages are a decreasing function of transport cost as well as whether these trade shocks were able to impact this relationship. Two distances are tested: to internal market; to external market. Outcomes suggest that distance to industry center (referred as internal market) is negatively associated to differences in wages. On the other hand, transport cost to foreign market is positively correlated to regional wages. These results suggest that firms might prefer to be closer to internal market and distant from external market.

In examining whether trade shocks have impacted these differences, outcomes show interesting results. First, those shocks have affected regional disparities in wages, but differently depending on which market. While domestic market has weakened its importance, external market has increased. It is safe to conclude therefore that economic shocks may affect regional disparities in wages. As it regards which shock is more effective in impacting these regional inequalities, results are not so obvious, since coefficients do not statistical difference. Trying to evaluate which market might be more affected, it is either not possible to conclude, because depending on which approach is used, different conclusions emerge.

In order to provide robustness check, further results are obtained by using market potential measures instead. By utilizing this other measure, demand linkages from the

internal market are confirmed as more influential in explaining regional disparities in wages than from foreign market. These results strongly illustrate the importance of the industry center to design the distribution of regional wage over the Brazilian landscape. For trade shock effect, some outcomes reinforced few previous results with transport cost measure but others continuous dubious.

One result strengthened by using market potential is that trade shock affects regional disparities. Now, not only external market becomes more relevant to explain wages differences across space, but also internal market has its force lessened after a trade shock. These results help to rule out the idea that a trade shock has no impact on regional inequalities, as found in Hanson (1997). On the other hand, some outcomes remain similar to previous findings. With regards to where trade shocks have impacted more, it is not feasible to observe since absolute values from internal and external market are not very distinguishable.

As it regards which shock is more powerful to affect these disparities, findings using market potential are pretty similar to those previous. In other words, it is difficult to establish a definitive conclusion. At first, results suggest that reducing the export cost is more relevant than decreasing import cost. However, these initial findings are not confirmed by estimating with fixed effects. Overall, it is safe to conclude that trade shocks may impact regional inequalities, but the strength of each shock as well as which market is more affected are still inconclusive.

Chapter 3 teaches us that a regional component is extremely important to explain differences in wages. Additionally, reducing cost to trade really do impact those disparities since after these two shocks regions close to external market become more relevant to explain them and internal market less. Extensions of this work might help highlight some

unanswered questions, such as examining which cost is more effective: export or import. Perhaps information at firm level may help tackle this issue more properly.

Government policies may impact firms' and individuals' location decisions. One way which might affect these decisions is by aiming to increase productivity, therefore growth. Chapter 4 investigates whether a government institution in Brazil (BNDES) was able to improve firms' productivity. Two channels were further explored: project size (small versus large projects); and location (poor versus rich regions). Descriptive statistics show that supported firms were twice more productive before receiving a loan than non-treated firms. After the loan, they become nearly three times more productive.

Investigating whether they are correlated to the government intervention, some positive association between firms' performance and BNDES loans is found only in most naive models. When more sophisticated models were used, then this association vanishes. Using a simple regression analysis, not a single distinction between size and location is found after controlling for all possible covariates, although some appear with few controls. Large projects in rich and poor areas do not present significant association. Controlling for unobservable characteristics of granted and non-granted firms, results evidence that only the performance of small projects in rich areas are positively correlated to the Bank's loans.

Results change considerably when controlling for all unobservable firms' attributes. Without controls, negative association between small projects' performance overall or only in poor areas with BNDES financial support is found. On the other hand, large projects in rich areas as well as in all of Brazil seem to be positively correlated to the Bank's loans. These results are not confirmed by using all covariates. The only significant outcome is a negative association between BNDES loans with projects developed in poor areas. When trying to investigate BNDES effects on firms' performance by estimating a dif-in-dif equation, no evidence is found, neither positive nor negative.

Further work should be pursued with better information with regards to the loans, such as amount, maturity period and so forth, in order to investigate whether government interventions are able to improve firms' productivity. Other measures of productivity, such as TFP, should also be used for a better understanding on how these policies affect firms' performance. Another channel to be explored is by introducing a theoretical model with firms' heterogeneity using Melitz (2003) framework for understanding how different firms react after being "treated".

Although there is certainly room for further research, some promising results have been produced in this study. Economic factors are relevant for individuals to migrate, but other factors seem to be associated as well. As results show, social amenities are associated to migrations flows in the Brazilian case, especially those negative. Due to the urbanization process worldwide, more people are living in cities than in rural areas nowadays. These results therefore might suggest that medium-sized cities might become more important in a near future.

For firms, economic factors are not only relevant but the main factor to explain firms' location decision. Chapter 3 teaches us that whenever economic drivers change, such as by trade shocks, importance of some regions raises and of others reduces. Overall, domestic markets are more influential in shaping regional wages, but trade shocks have made regions close to external market more attractive for firms. In a world of globalization, where trade barriers tend to reduce, it is important to understand the geographical changes towards external markets. Regions close to frontiers and coasts may become much more developed than in countryside areas.

Whether government interventions are able to lead growth by improving firms' productivity, it was not possible to find a conclusive answer. Since results of Chapter 4 fail to reject the hypothesis of no effect on firms' productivity by government support, lessons

left are inconclusive. Either BNDES loans are not able to impact firms' productivity, or methods implemented are not able to detect during this particular period.

Nevertheless, these are fairly important and useful set of results for agents' decision related to location and government interventions, which have been produced in this thesis. Those were done by filling in some gaps in the existing literature. This should be of interest for the assessment of the place of this regional economic class.

Appendices

Appendix 2.I: Variable List of Chapter 2

1.a Information about AMC

The number of Brazilian municipalities has increased from 3,951 in 1970 to 5,507 in 2000, where most of this change occurred during the 1990's. This huge increase in the number of cities hampers the possibility of research for long periods. In order to overcome this issue, IPEA has aggregated actual municipalities to what it was before and creating therefore a new classification called Minimum Comparable Areas (*Areas Minimas Comparaveis* - AMC). Using 2000 and 1970 as the benchmark years, there are 3,659 cities which information is available for this study. More information can be accessed at IPEADATA website (www.ipeadata.gov.br).

1.b Detailed Description of variables used in this paper.

1) **Migration flows**: Total number of people who were living in an urban area of a particular city at Census year, but 5 years before were living in another urban area at a different city. Any migration from and/or to rural areas are not included in the total migration flows. The Brazilian Census is elaborated by the Brazilian Statistical Institute (*Instituto Brasileiro de Geografia e Estatistica – IBGE*) and the available data are from 1991 and 2000.

2) **Violence**: Measures are provided by SIM-DATASUS from the Health Ministry which includes homicides per 100 thousand habitants from 1980 until 2002. There is also information on suicide and car victims per 100 thousand habitants for the same period.

3) **Security Measure**: The amounts of expenditure in security and in total were extracted from the state government budget from 1985 until 2000. The average of 5 years between the periods of change of dwellings from migrants is used as instrument for violence. Source is National Treasury Secretariat from the Finance Ministry.

4) **Distance**: The distance between origin and destination of migration flows are calculated by Great Circle Formula using the centroids of cities as a fixed point. Distance to the state capital and Brasilia (Brazil's capital) are available from IPEADATA website.

5) **Nominal Wages**: Total Wages and number of employees from 1985 come from the Economic Census from that particular year. This Economic Census includes Manufacturing Census, Services Census and Commerce Census. From 1996, data are from Firms Central Register (*Cadastro Central de Empresas – CEMPRE*), which has information about wages and number of employees of any Brazilian firm. Services and Commerce were treated jointly, since they present similar structures.

6) **Cost of Living**: Capital Residential Stock created by Applied Economical Research Institute (Instituto de Pesquisa Economica Aplicada - IPEA) for the Census years (1980, 1991 and 2000). Value for 1985 and 1996 are created by interpolating the figures from the census. The number of houses comes from the IBGE census and the years 1985 and 1996 are also created by interpolating census figures.

7) **GDP** measures source is IPEA. They are used for creating the Market Potential measure; IPEA has information for the years 1985 and 1996 which is calculated base on the economical censuses and similar publications.

8) **Share of Urban Population:** Census data from 1980 to 2000 is used to create the percentage of urban population at each city. Numbers for 5 years before the census are interpolated by Censuses figures. Correlation between these created measures and the actual value on 1991 and 2000 (census years) are extremely high. Both used and no substantial difference is found.

9) **Unemployment Rate**: Using information on how many employees in agriculture, manufacturing and services together with information on how many active economic people, it was possible to create an unemployment rate at each location. First, workers in all sectors are added to create how many people were working at each location. After, the ratio of economic active people over total number of employees is used to create an unemployment rate measure. The sources of sectors employment are the Censuses and related publications from IBGE. The source of Active Economical Population is a measure from IPEA based on Census data. All data referred to 5 years before the census.

10) **Housing Infrastructure**: The percentage of housing having electricity, water and sewage was calculated by using the number of residences from the census data, elaborated by the Brazilian Statistical Institute (*IBGE*).

11) **Education**: There are two measures of education used here. The first measure is the number of years spent on schooling for any individual above 25 years old. The second is the percentage of illiterate people older than 24 years. Both measures are available at Brazilian Census from 1991 and 2000.

12) **Money**: The total of banking deposit and savings at each location is used as a measure for the Informal Sector. The source of them is the Brazilian Central Bank.

13) **Share of Male**: The percentage of male in urban population is created by using information from the Brazilian Censuses from 1991 and 2000.

14) **Health Indicators**: There are two measures. One is the number of doctors per one thousand habitants and the other is the percentage of nurses with at least an undergraduate degree. Both were elaborated by UNDP using Brazilian Census information.

15) **Sector's Share**: IPEA elaborated both GDP, manufacturing and services, for the period analyzed in this paper. They are used to create the share of each sector on every municipality GDP.

16) **Negative Amenity**: Population density was calculated by using the total population of a particular city and its area. The number of people is originated from Census and area from IPEADATA website.

17) **Positive Amenity**: Number of firms in the service sector at each city, which the source is IBGE.

18) **Export Status**: There is information on export volume at each location for 2003/04. A dummy for export status was created by assuming that a city which has exported more than a certain amount was considered an export region ever. Three thresholds are used: US$ 1 billion; US$ 500 billion; and US$ 1 trillion. All of them showed similar results on the estimation. Source is International Trade Secretary (*Secretaria de Comercio Exterior – SECEX*);

Appendix 2.II: Further Results of Chapter 2

Table 2.II.a show results using level instead of taking logarithm of each variable. Column (i) show OLS result using only distance and population share as explanatory variables. Column (ii) presents OLS using all covariates. 2SLS estimation results using only security measure as instruments are presented in column (iii) and Gini index added as instrument in column (iv). Last column present results using destination FE.

Table 2.II.b show results using a random sample of 500 thousand migration flows, including where no migration record between two localities existed. First column present OLS results, followed by 2SLS using only Security measure. Gini inequality index is added as instrument at results of column (iii). Column (iv) present Poison result using only non-zero migration flows, while column (v) using all migration flows. Last column reports zero inflated method.

Table 2.II.c report full results of the traditional equation model. First two columns show OLS results, using only distance and population share for the first and all covariates for the second. Column (iii) present outcomes using security measure as instrument in a 2SLS. The last two columns report Poison results, either by using only distance and population share (column iv), or by utilizing all controls (column v).

Table 2.II.a: Results with Levels instead of Logs

	OLS	OLS	2SLS	2SLS	FE
Distance	0	0	0	0	0
	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***
Share of Urban Pop.	0.935	-28.01	-49.37	-34.116	-18.725
	(0.194)***	(12.254)**	(12.917)***	(12.428)***	(8.878)**
Manuf. Real Wages		0	0	0	0
		(0.0)	(0.0)	(0.0)	(0.0)
Service Real Wages		0	0.001	0.001	0
		(0.0)	(0.000)**	(0.000)*	(0.0)
Market Potential		-0.893	2.018	-0.058	-1.735
		(0.385)**	(1.174)*	-0.658	(0.333)***
Unemployment Rate		0.002	0.003	0.002	-0.001
		(0.001)***	(0.001)***	(0.001)***	(0.0)
Violence		-0.001	-0.021	-0.007	0
		(0.0)	(0.005)***	(0.003)***	(0.0)
Electricity		-0.005	0.001	-0.004	0.005
		(0.003)*	(0.0)	(0.0)	(0.003)*
Water Supply		0.002	-0.003	0	-0.004
		(0.0)	(0.0)	(0.0)	(0.002)**
Sewage		0	-0.003	-0.001	0.001
		(0.0)	(0.001)**	(0.0)	(0.0)
GDP Growth		0	0	0	0
		(0.0)	(0.0)	(0.0)	(0.0)
Education		0.034	0.055	0.045	0.01
		(0.007)***	(0.011)***	(0.008)***	(0.005)**
Illiteracy		0.008	0.014	0.011	-0.002
		(0.002)***	(0.002)***	(0.002)***	-0.002
Informal Sector		-0.475	-0.425	-0.464	-0.298
		(0.160)***	(0.183)**	(0.165)***	(0.063)***
% of Male		27.593	48.354	33.499	18.456
		(11.961)**	(12.571)***	(12.121)***	(8.983)**
Dist. State Capital		0	0	0	-
		(0.0)	(0.0)	(0.0)	-
Dist. Brasilia		0	0	0	-
		(0.000)*	(0.000)**	(0.000)**	-
Doctors		-0.001	0	0	-0.001
		(0.0)	(0.0)	(0.0)	(0.000)*
Skilled Nurses		0	0	0	0
		(0.0)	(0.0)	(0.0)	(0.0)
Share of Manuf.		0.016	0.017	0.016	0.001
		(0.002)***	(0.004)***	(0.003)***	(0.0)
Share of Services		0.01	0.016	0.012	0
		(0.003)***	(0.005)***	(0.003)***	(0.0)
Population Density		0.269	0.465	0.336	-0.093
		(0.054)***	(0.066)***	(0.055)***	(0.042)**
Population Density Square		-0.141	-0.173	-0.152	0.022
		(0.023)***	(0.033)***	(0.024)***	(0.074)
Entertainment		1.543	0.122	1.083	0.101
		(0.666)**	(0.8)	(0.707)	(0.262)
Observations	275,308	275,278	275,278	275,278	275,278

Table 2.II.b: Results with a Random Sample of 500,000 Migration Flows

	(i) OLS	(ii) 2SLS	(iii) 2SLS	(iv) Poisson shmig>0	(v) Poisson all	(vi) ZIM
Violence	0.00	-0.00	-0.00	0.09	0.14	0.14
	(0.00)	(0.000)***	(0.000)*	(0.11)	(0.15)	(0.31)
Distance	0.00	0.00	0.00	-0.62	-1.57	-1.56
	(0.000)***	(0.000)***	(0.000)***	(0.023)***	(0.030)***	(0.059)***
Share of Urban Pop.	-2.03	-3.00	-2.20	-692.01	-2400.05	-2438.91
	(0.380)***	(0.406)***	(0.387)***	(379.727)*	(450.969)***	(811.937)***
Manuf. Real Wages	0.00	0.00	0.00	3.15	-1.92	-1.11
	(0.00)	(0.00)	(0.00)	(11.88)	(4.75)	(31.11)
Service Real Wages	0.01	0.01	0.01	37.36	36.71	36.70
	(0.003)***	(0.003)***	(0.003)***	(17.028)**	(1.896)***	(19.108)*
Unemployment Rate	0.00	0.00	0.00	0.17	0.44	0.43
	(0.00)	(0.000)***	(0.00)	(0.077)**	(0.085)***	(0.223)*
Market Potential	-0.08	-0.02	-0.07	-71.47	-148.21	-148.33
	(0.010)***	(0.013)*	(0.011)***	(15.103)***	(31.023)***	(55.076)***
GDP Growth	0.00	0.00	0.00	-0.46	0.44	0.59
	(0.00)	(0.000)*	(0.00)	(1.16)	(1.12)	(2.88)
Dist. to State Capital	0.00	0.00	0.00	-0.08	-0.12	-0.12
	(0.000)***	(0.000)***	(0.000)***	(0.028)***	(0.040)***	(0.071)*
Distance to Brasilia	0.00	0.00	0.00	-0.08	-0.20	-0.19
	(0.000)*	(0.000)**	(0.000)*	(0.05)	(0.063)***	(0.13)
Electricity	0.00	0.00	0.00	-0.95	-2.44	-2.50
	(0.000)***	(0.000)***	(0.000)***	(0.61)	(0.591)***	(1.94)
Water Supply	0.00	0.00	0.00	0.31	1.04	1.08
	(0.00)	(0.000)***	(0.00)	(0.33)	(0.440)**	(1.05)
Sewage	0.00	0.00	0.00	0.00	-0.25	-0.25
	(0.000)**	(0.000)***	(0.000)***	(0.11)	(0.120)**	(0.28)
Years of Schooling	0.00	0.00	0.00	2.59	4.03	4.09
	(0.000)***	(0.000)***	(0.000)***	(0.720)***	(0.936)***	(2.178)*
Illiteracy	0.00	0.00	0.00	0.89	0.64	0.64
	(0.000)***	(0.000)***	(0.000)***	(0.253)***	(0.341)*	(0.83)
Informal Sector	-0.03	-0.03	-0.03	-12.69	-15.88	-15.54
	(0.007)***	(0.007)***	(0.007)***	(6.735)*	(9.386)*	(16.98)
% of Male	2.02	2.97	2.18	686.53	2336.73	2374.03
	(0.381)***	(0.406)***	(0.388)***	(369.816)*	(438.279)***	(791.651)***
Doctors	0.00	0.00	0.00	0.16	0.42	0.42
	(0.000)*	(0.000)***	(0.000)**	(0.10)	(0.112)***	(0.27)
Nurses	0.00	0.00	0.00	0.11	0.35	0.35
	(0.00)	(0.00)	(0.00)	(0.059)*	(0.070)***	(0.161)**
Share of Manuf.	0.00	0.00	0.00	1.35	2.38	2.41
	(0.00)	(0.000)***	(0.00)	(0.310)***	(0.244)***	(0.715)***
Share of Services	0.00	0.00	0.00	0.75	1.55	1.50
	(0.00)	(0.00)	(0.00)	(0.427)*	(0.312)***	(0.860)*
Negative Amenities	0.05	0.06	0.05	7.37	23.00	23.60
	(0.003)***	(0.003)***	(0.003)***	(2.903)**	(3.359)***	(6.130)***
Positive Amenities	0.11	0.10	0.11	4.42	-23.70	-26.15
	(0.029)***	(0.029)***	(0.029)***	(26.31)	(33.78)	(61.49)
Export Status	0.00	0.00	0.00	-0.07	0.17	-
	(0.000)***	(0.000)***	(0.000)***	(0.12)	(0.15)	-
Observations	499,304	499,304	499,304	5,144	499,304	499,304

Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%

Table 2.II.c: Traditional Gravity Equation

Dep. Variable: Share of Migrants	OLS	OLS	2SLS	FE	Poisson
Distance	-0.317	-0.417	-0.413	-	-0.889
	(0.002)***	(0.003)***	(0.004)***	-	(0.014)***
Share of Pop	0.296	-2.622	-2.911	-1.388	-11.495
	(0.001)***	(0.185)***	(0.212)***	(1.307)	(0.625)***
Share of Pop Origin	0.213	1.273	1.308	4.395	4.153
	(0.002)***	(0.183)***	(0.197)***	(1.253)***	(0.717)***
Violence		0.011	-0.085	-0.037	0.004
		(0.004)***	(0.028)***	(0.020)*	(0.011)
Violence Origin		0.062	0.053	0.052	0.097
		(0.004)***	(0.020)***	(0.019)***	(0.012)***
Population Density		-0.025	-0.076	-0.642	0.026
		(0.025)	(0.028)***	(0.160)***	(0.083)
Population Density Square		-0.011	-0.012	-0.024	-0.007
		(0.001)***	(0.001)***	(0.017)	(0.004)*
Entertainment		0.11	0.111	0.024	0.324
		(0.010)***	(0.010)***	(0.051)	(0.039)***
Population Density Origin		0.223	0.219	0.35	0.306
		(0.028)***	(0.028)***	(0.187)*	(0.086)***
Population Density Sq Origin		0.016	0.016	0.03	0.039
		(0.002)***	(0.002)***	(0.019)	(0.006)***
Entertainment Origin.		-0.039	-0.042	0.08	0
		(0.010)***	(0.010)***	(0.05)	(0.03)
Manuf. Real Wage		-0.032	-0.019	0.04	-0.045
		(0.006)***	(0.007)***	-0.028	(0.018)**
Service Real Wage		-0.001	0.035	0.023	-0.004
		(0.008)	(0.012)***	(0.032)	(0.025)
Unemployment Rate	0.048	0.072	0	0.191	0.048
	(0.010)***	(0.012)***	(0.05)	(0.043)***	(0.010)***
Market Potential	-0.156	-0.125	1.435	-0.351	-0.156
	(0.006)***	(0.010)***	(0.306)***	(0.022)***	(0.006)***
GDP Growth	0.012	0.01	0	-0.004	0.012
	(0.003)***	(0.003)***	(0.007)	-0.016	(0.003)***
Distance to State Capital	-0.007	-0.017	0	-0.058	-0.007
	(0.003)**	(0.004)***	(0.0)	(0.011)***	(0.003)**
Distance to Brasilia	-0.019	-0.013	0	-0.034	-0.019
	(0.005)***	(0.005)**	(0.0)	(0.010)***	(0.005)***
Electricity	-0.331	-0.274	0.009	-1.032	-0.331
	(0.043)***	(0.045)***	(0.202)	(0.144)***	(0.043)***
Water Supply	0.019	-0.021	-0.183	0.002	0.019
	(0.015)	(0.019)	(0.095)*	(0.057)	(0.015)
Sewage	-0.01	-0.009	-0.011	-0.004	-0.01
	(0.002)***	(0.003)***	(0.009)	(0.008)	(0.002)***
Educational Level	0.101	0.102	-0.028	0.698	0.101
	(0.028)***	(0.028)***	(0.246)	(0.086)***	(0.028)***
Informal Sector	0.014	0.007	0.024	-0.067	0.014
	(0.005)***	(0.005)	(0.014)*	(0.015)***	(0.005)***
Share of Manuf Sector	0.042	0.027	-0.006	0.014	0.042
	(0.004)***	(0.006)***	(0.022)	(0.013)	(0.004)***
Share of Service Sector	0.013	0.008	-0.059	-0.096	0.013
	(0.011)	(0.011)	(0.048)	(0.032)***	(0.011)
Manuf Real Wage Origin	0.03	0.033	-0.072	0.043	0.03
	(0.005)***	(0.006)***	(0.026)***	(0.015)***	(0.005)***
Service Real Wage Origin	0.026	0.032	0.058	0.102	0.026
	(0.007)***	(0.010)***	(0.032)*	(0.024)***	(0.007)***
Market Potential Origin	-0.166	-0.164	0.566	-0.541	-0.166
	(0.008)***	(0.010)***	(0.39)	(0.036)***	(0.008)***
Unemployment Rate Origin	-0.043	-0.04	0.036	-0.083	-0.043
	(0.010)***	(0.011)***	(0.05)	(0.038)**	(0.010)***

Continuing Table 2.II.c

GDP Growth Origin	-0.006	-0.006	0.012	-0.012	-0.006
	(0.003)*	(0.003)*	(0.007)*	(0.013)	(0.003)*
Distance to State Capital Origin	0.026	0.025	0	0.055	0.026
	(0.005)***	(0.006)***	0	(0.021)***	(0.005)***
Distance to Brasilia Origin	0.01	0.018	0	-0.026	0.01
	-0.008	(0.009)*	0	(0.021)	(0.008)
Electricity Origin	-0.257	-0.216	-0.651	-0.925	-0.257
	(0.038)***	(0.040)***	(0.178)***	(0.103)***	(0.038)***
Water Supply Origin	0.078	0.074	0.044	0.106	0.078
	(0.015)***	(0.016)***	(0.094)	(0.048)**	(0.015)***
Sewage Origin	-0.019	-0.017	-0.018	-0.02	-0.019
	(0.002)***	(0.002)***	(0.008)**	(0.006)***	(0.002)***
Educational Level Origin	-0.107	-0.12	-0.009	-0.052	-0.107
	(0.027)***	(0.027)***	(0.231)	(0.096)	(0.027)***
Informal Sector Origin	-0.009	-0.009	-0.011	-0.021	-0.009
	(0.005)*	(0.005)*	(0.013)	(0.02)	(0.005)*
Share of Male Origin	-1.026	-1.052	-3.882	-3.489	-1.026
	(0.179)***	(0.197)***	(1.240)***	(0.695)***	(0.179)***
Nurses Origin	-0.001	0	-0.002	0.008	-0.001
	(0.001)	(0.001)	(0.005)	(0.003)**	(0.001)
Share of Manuf Origin	-0.033	-0.038	0.014	-0.111	-0.033
	(0.004)***	(0.005)***	(0.021)	(0.014)***	(0.004)***
Share of Services Origin	-0.029	-0.028	0.039	-0.08	-0.029
	(0.011)***	(0.011)**	(0.048)	(0.032)**	(0.011)***
Constant	-7.038	-8.229	-7.935	5.818	-3.723
	(0.031)***	(0.126)***	(0.139)***	(2.823)**	(0.436)***
Observations	275,308	113,307	113,307	275,308	113,307
R-squared	0.26	0.28	0.27	0.03	0.06

Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%

The next two tables show results of restricted model and traditional model using contiguity spatial matrix (Table 2.II.d) and nearest spatial matrix (Table 2.II.e). In both tables, columns follow the same order. The first four columns report outcomes using the restricted model; while the last four, using traditional gravity equation. First column of each model show OLS results, followed by 2SLS using security measure as instrument. Third column report FE outcomes and last Poisson.

Table 2.II.d: Results Using Contiguity Matrix

Share of Migrants	OLS	2SLS	FE Dest	Poisson	OLS	2SLS	FE Pairw	Poisson
Violence	0.056	0.657	-0.007	-0.039	0.006	-0.394	-0.044	-0.005
	(0.023)**	(0.816)	(0.027)	(0.03)	(0.006)	(0.228)*	(0.024)*	(0.015)
Spatial Viol Cont	-0.03	-0.906	0.006	-0.066	0.001	0.238	-0.005	-0.009
	(0.021)	(0.65)	(0.027)	(0.030)**	(0.006)	(0.167)	(0.026)	(0.016)
Distance	-0.74	-0.723	-0.776	-0.619	-0.414	-0.397	-	-0.884
	(0.032)***	(0.034)***	(0.035)***	(0.061)***	(0.004)***	(0.008)***	-	(0.014)***
Share of Pop	-6.57	-11.567	0.751	-2.417	-2.594	-1.92	-2.369	-11.399
	(0.891)***	(2.864)***	(2.227)	(1.216)**	(0.196)***	(0.551)***	(1.387)*	(0.645)***
Manuf. Real Wage	-0.095	-0.018	0.076	-0.141	-0.029	-0.005	0.047	-0.037
	(0.027)***	(0.069)	(0.036)**	(0.039)***	(0.006)***	(0.011)	(0.029)	(0.019)**
Service Real Wage	0.038	0.058	-0.001	0.086	-0.004	0.076	0.028	0.008
	(0.033)	(0.168)	(0.033)	(0.043)**	(0.008)	(0.037)**	(0.033)	(0.026)
Unemployment Rate	0.024	0.025	-0.062	0.126	0.044	0.109	-0.021	0.175
	(0.048)	(0.141)	(0.067)	(0.062)**	(0.010)***	(0.035)***	(0.052)	(0.042)***
Market Potential	-0.107	0.164	0.864	-0.315	-0.151	-0.12	1.296	-0.333
	(0.033)***	(0.116)	(0.279)***	(0.056)***	(0.006)***	(0.016)***	(0.319)***	(0.023)***
GDP Growth	0.007	0.015	0.006	0.006	0.013	0.007	0.001	0.004
	(0.01)	(0.024)	(0.009)	(0.011)	(0.003)***	(0.006)	(0.008)	(0.015)
Distance to State Capital	-0.04	-0.103	-	-0.055	-0.007	-0.015	-	-0.068
	(0.013)***	(0.033)***	-	(0.014)***	(0.003)**	(0.005)***	-	(0.012)***
Distance to Brasilia	-0.071	0.003	-	-0.039	-0.017	-0.029	-	-0.03
	(0.023)***	(0.051)	-	(0.020)**	(0.005)***	(0.012)**	-	(0.010)***
Electricity	-0.909	-0.801	0.22	-0.819	-0.356	-0.294	0.187	-1.052
	(0.217)***	(0.371)**	(0.337)	(0.257)***	(0.044)***	(0.056)***	(0.216)	(0.146)***
Water Supply	0.02	-0.15	-0.286	-0.078	0.012	-0.024	-0.284	-0.023
	(0.079)	(0.159)	(0.104)***	(0.088)	(0.016)	(0.023)	(0.106)***	(0.058)
Sewage	-0.031	-0.044	0.003	0.001	-0.01	0.002	-0.012	-0.003
	(0.009)***	(0.017)***	(0.009)	(0.01)	(0.003)***	(0.008)	(0.009)	(0.008)
Educational Level	0.167	0.507	0.377	0.343	0.093	0.036	-0.222	0.687
	(0.163)	(0.497)	(0.349)	(0.212)	(0.029)***	(0.048)	(0.263)	(0.085)***
Illiteracy	0.128	0.207	-0.154	0.149				
	(0.060)**	(0.315)	(0.166)	(0.071)**				
Informal Sector	0.039	0.002	-0.004	-0.005	0.019	0.014	0.016	-0.057
	(0.016)**	(0.032)	(0.023)	(0.02)	(0.005)***	(0.007)*	-0.014	(0.015)***
Share of Male	7.013	11.869	-0.587	2.953	2.738	2.21	2.758	11.667
	(0.853)***	(2.789)***	(2.208)	(1.196)**	(0.191)***	(0.502)***	(1.370)**	(0.642)***
Nurses	-0.002	0.005	-0.001	-0.003	-0.017	-0.016	0.006	0.005
	(0.005)	(0.009)	(0.006)	(0.006)	(0.001)***	(0.002)***	-0.006	-0.005
Share of Manuf Sector	0.059	-0.022	-0.062	0.09	0.04	0.014	-0.003	0.006
	(0.018)***	(0.051)	(0.031)**	(0.024)***	(0.005)***	(0.01)	(0.024)	(0.013)
Share of Service Sector	-0.002	-0.03	0.032	-0.165	0.012	0.018	-0.053	-0.108
	(0.044)	(0.07)	(0.05)	(0.047)***	(0.012)	(0.015)	(0.051)	(0.037)***
Population Density	0.105	0.289	-0.422	-0.114	-0.042	-0.174	-0.704	-0.044
	(0.122)	(0.341)	(0.201)**	(0.156)	(0.026)	(0.064)***	(0.170)***	(0.081)
Population Density Sq.	0.005	0.006	-0.041	0.018	-0.011	-0.015	-0.039	-0.009
	(0.004)	(0.019)	(0.020)**	(0.005)***	(0.001)***	(0.002)***	(0.018)**	(0.004)**
Entertainment	0.002	-0.03	0.074	0.136	0.107	0.1	-0.01	0.322
	(0.043)	(0.127)	(0.059)	(0.055)**	(0.011)***	(0.016)***	(0.054)	(0.039)***

Continuing Table 2.II.d

Violence Origin					0.038 (0.006)***	-0.969 (0.572)*	0.021 (0.022)	0.059 (0.018)***
Spatial Violence Origin					0.034 (0.006)***	0.969 (0.558)*	0.066 (0.024)***	0.062 (0.019)***
Manuf Real Wage Origin					0.026 (0.006)***	-0.013 (0.028)	-0.089 (0.027)***	0.038 (0.016)**
Service Real Wage Origin					0.024 (0.008)***	0.143 (0.061)**	0.059 (0.034)*	0.081 (0.022)***
Market Potential Origin					-0.17 (0.009)***	-0.235 (0.048)***	0.465 (0.405)	-0.547 (0.037)***
Share of Pop Origin					1.396 (0.196)***	4.416 (1.837)**	4.273 (1.320)***	3.874 (0.695)***
Unemployment Rate Origin					-0.04 (0.011)***	0.081 (0.066)	0.005 (0.052)	-0.083 (0.040)**
GDP Growth Origin					-0.008 (0.003)**	-0.012 (0.005)**	0.015 (0.008)**	-0.007 (0.011)
Distance to State Capital Origin					0.029 (0.005)***	0.129 (0.063)**	- -	0.074 (0.021)***
Distance to Brasilia Origin					0.002 (0.009)	0.052 (0.022)**	- -	-0.041 (0.022)*
Electricity Origin					-0.26 (0.039)***	-0.071 (0.088)	-0.665 (0.184)***	-0.857 (0.105)***
Water Supply Origin					0.084 (0.016)***	0.145 (0.052)***	0.099 (0.098)	0.128 (0.051)**
Sewage Origin					-0.02 (0.002)***	0.002 (0.011)	-0.017 (0.009)**	-0.025 (0.006)***
Educational Level Origin					-0.136 (0.028)***	-0.433 (0.173)**	-0.11 (0.244)	-0.199 (0.093)**
Informal Sector Origin					-0.007 (0.005)	0.005 (0.011)	-0.007 (0.014)	-0.006 (0.021)
Share of Male Origin					-1.154 (0.191)***	-3.962 (1.729)**	-3.671 (1.305)***	-3.23 (0.676)***
Nurses Origin					-0.001 (0.001)	-0.013 (0.008)*	-0.001 (0.006)	0.006 (0.003)
Share of Manuf Origin					-0.026 (0.005)***	-0.004 (0.019)	0.014 (0.022)	-0.093 (0.013)***
Share of Services Origin					-0.037 (0.012)***	0.028 (0.042)	0.07 (0.05)	-0.098 (0.032)***
Population Density Origin					0.223 (0.029)***	0.379 (0.103)***	0.379 (0.198)*	0.306 (0.086)***
Population Density Sq Origin					0.016 (0.002)***	0.048 (0.019)**	0.038 (0.020)*	0.037 (0.006)***
Entertainment Origin					-0.04 (0.011)***	-0.001 (0.03)	0.062 (0.053)	0.001 (0.032)
Observations	141727	141727	141727	141727	101270	101270	101270	101270
R-squared	0.32		0.34		0.28		0.03	

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 2.II.e: Four Nearest Neighbours Matrix

Dep. Variable: Share Migrants	OLS	2SLS	FE Dest.	Poisson	OLS	2SLS	FE Pairw	Poisson
Violence	0.085	6.503	-0.003	0.04	0.009	1.558	-0.044	0.028
	(0.024)***	(17.689)	(0.026)	(0.029)	(0.005)*	(2.978)	(0.022)**	(0.019)
Spatial Violence Near.	-0.06	-4.369	0.014	-0.145	-0.001	-1.022	0.004	-0.044
	(0.022)***	(10.54)	(0.026)	(0.029)***	(0.006)	(1.854)	(0.024)	(0.020)**
Distance	-0.738	-0.838	-0.771	-0.622	-0.417	-0.463	-	-0.89
	(0.031)***	(0.341)**	(0.033)***	(0.058)***	(0.003)***	(0.097)***	-	(0.014)***
Pop. Share	-6.656	2.495	1.216	-2.824	-2.589	-2.68	-1.284	-11.656
	(0.839)***	(40.423)	(2.147)	(1.119)**	(0.185)***	(0.513)***	(1.318)	(0.633)***
Real Manuf. Wage	-0.083	-0.023	0.07	-0.128	-0.03	-0.085	0.045	-0.04
	(0.026)***	(0.27)	(0.035)**	(0.038)***	(0.006)***	(0.125)	(0.028)	(0.018)**
Real Service Wage	0.027	-1.623	-0.008	0.058	-0.002	-0.308	0.02	-0.006
	(0.032)	(4.932)	(0.033)	(0.041)	(0.008)	(0.621)	(0.032)	(0.025)
Unemployment Rate	0.03	-1.079	-0.056	0.124	0.051	-0.182	0.006	0.194
	(0.046)	(3.29)	(0.066)	(0.060)**	(0.010)***	(0.465)	(0.05)	(0.043)***
Mkt Potential	-0.095	0.064	0.943	-0.284	-0.155	-0.064	1.452	-0.329
	(0.032)***	(0.795)	(0.279)***	(0.052)***	(0.006)***	(0.11)	(0.307)***	(0.022)***
Growth	0.005	0.251	0.005	0.004	0.012	0.059	0	-0.004
	(0.009)	(0.715)	(0.009)	(0.0110	(0.003)***	(0.087)	(0.007)	(0.016)
Distance Brasilia	-0.039	0.052	-	-0.053	-0.007	0.016	-	-0.062
	(0.012)***	(0.487)	-	(0.013)***	(0.003)**	(0.06)	-	(0.011)***
Distance State Capital	-0.078	-0.39	-	-0.051	-0.019	-0.071	-	-0.034
	(0.021)***	(0.952)	-	(0.019)***	(0.005)***	(0.106)	-	(0.010)***
Electricity	-0.884	-1.896	0.248	-0.743	-0.339	-0.572	-0.016	-1.048
	(0.205)***	(3.571)	(0.322)	(0.239)***	(0.043)***	(0.533)	(0.202)	(0.143)***
Water Supply	0.023	0.812	-0.316	-0.072	0.018	0.184	-0.197	-0.006
	(0.076)	(2.698)	(0.097)***	(0.083)	(0.015)	(0.369)	(0.095)**	(0.057)
Sewage	-0.035	-0.203	0.006	-0.004	-0.01	-0.063	-0.011	-0.005
	(0.008)***	(0.435)	(0.009)	(0.01)	(0.002)***	(0.098)	(0.009)	(0.008)
Education	0.146	-3.737	0.45	0.236	0.105	0.166	0.027	0.714
	(0.155)	(12.444)	(0.333)	(0.208)	(0.028)***	(0.139)	(0.249)	(0.086)***
Informal Sector	0.033	0.142	-0.009	0	0.014	0.057	0.023	-0.069
	(0.015)**	(0.432)	(0.022)	(0.019)	(0.005)***	(0.09)	(0.014)*	(0.015)***
Share of Male	7.097	-2.767	-1.041	3.328	2.714	2.32	1.664	11.883
	(0.803)***	(42.407)	(2.124)	(1.101)***	(0.180)***	(1.297)*	(1.302)	(0.627)***
Qualified Nurses	-0.003	-0.058	0	-0.005	-0.018	-0.035	0.007	0.004
	-0.004	(0.138)	(0.006)	(0.006)	(0.001)***	(0.032)	(0.006)	(0.004)
Manuf. Share	0.052	-0.011	-0.052	0.081	0.041	0.094	-0.013	0.009
	(0.017)***	(0.257)	(0.029)*	(0.024)***	(0.004)***	(0.123)	(0.022)	(0.013)
Service Share	-0.002	-0.107	0.037	-0.167	0.011	-0.013	-0.077	-0.101
	(0.04)	(0.506)	(0.049)	(0.043)***	(0.011)	(0.044)	(0.048)	(0.032)***
Population Density	0.084	2.134	-0.535	-0.096	-0.025	0.397	-0.61	0.036
	(0.125)	(5.847)	(0.198)***	(0.154)	(0.025)	(0.864)	(0.162)***	(0.084)
Population Density Sq.	0.004	0.159	-0.053	0.018	-0.011	0.001	-0.019	-0.007
	(0.006)	(0.457)	(0.019)***	(0.006)***	(0.001)***	(0.025)	(0.017)	(0.004)*
Entertainment	0.021	-1.003	0.093	0.153	0.11	0.07	0.032	0.322
	(0.044)	(2.906)	(0.057)	(0.054)***	(0.010)***	(0.076)	(0.051)	(0.039)***

Continuing Table 2.II.e

Violence Origin					0.036	1.671	0.034	0.057
					(0.005)***	(3.394)	(0.021)	(0.020)***
Spatial Viol. Near. Origin					0.039	-1.658	0.037	0.066
					(0.006)***	(3.415)	(0.022)*	(0.021)***
Real Manuf. Wage Orig.					0.027	0.104	-0.075	0.037
					(0.005)***	(0.145)	(0.026)***	(0.015)**
Real Service Wage Orig.					0.027	-0.139	0.066	0.104
					(0.007)***	(0.379)	(0.032)**	(0.024)***
Mkt Potential Origin					-0.174	0.002	0.617	-0.559
					(0.008)***	(0.334)	(0.392)	(0.036)***
Share Pop. Origin					1.408	-3.784	4.584	4.464
					(0.184)***	(10.11)	(1.262)***	(0.729)***
Unemployment Rate Origin					-0.043	-0.211	0.035	-0.081
					(0.010)***	(0.369)	(0.05)	(0.037)**
Growth Origin					-0.007	0.011	0.014	-0.012
					(0.003)**	(0.035)	(0.007)*	(0.013)
Dist. Brasilia Origin					0.034	-0.238	-	0.068
					(0.005)***	(0.532)	-	(0.020)***
Dist. State Origin					0.006	0.037	-	-0.031
					(0.008)	(0.035)	-	(0.021)
Electricity Origin					-0.26	-0.683	-0.68	-0.91
					(0.038)***	(0.948)	(0.179)***	(0.103)***
Water Supply Origin					0.084	-0.037	0.071	0.118
					(0.015)***	(0.21)	(0.094)	(0.048)**
Sewage Origin					-0.019	-0.047	-0.018	-0.02
					(0.002)***	(0.063)	(0.009)**	(0.006)***
Education Origin					-0.116	0.467	-0.053	-0.076
					(0.027)***	(1.19)	(0.233)	(0.094)
Informal Sector Origin					-0.008	-0.038	-0.011	-0.018
					(0.005)	(0.053)	(0.014)	(0.02)
Male Share Origin					-1.163	3.809	-4.065	-3.802
					(0.180)***	(9.616)	(1.248)***	(0.706)***
Doctor Origin					-0.001	0.004	-0.001	0.008
					(0.001)	(0.009)	(0.005)	(0.003)**
Manuf. Share Origin					-0.029	-0.123	0.012	-0.104
					(0.004)***	(0.173)	(0.021)	(0.014)***
Service Share Origin					-0.025	-0.2	0.042	-0.076
					(0.011)**	(0.341)	(0.048)	(0.032)**
Population Density Origin					0.235	-0.259	0.359	0.322
					(0.028)***	(0.99)	(0.189)*	(0.085)***
Population Density Sq. Origin					0.017	-0.043	0.03	0.041
					(0.002)***	(0.126)	(0.02)	(0.006)***
Entertainment Origin					-0.038	-0.071	0.076	0.003
					(0.010)***	(0.09)	(0.051)	-0.03
Observations	150483	150483	150483	150483	112661	112661	112661	112661
R-squared	0.32		0.34		0.28		0.03	

Annex 3.I: Further Results of Chapter 3

Tables 3.I.a shows results using distance as explanatory variable but not instrumenting human capital. The table has the same structure as Table 3.3, which is the following:

The first six columns (i to vi) presents results with time-invariant distance to industry centre, while the last four columns (vii to x) shows using the distance to the industry centre weighted by its importance. The first column of each method (i, iv and ix) presents results with no trade effect. Columns ii, v, vii and x use state government expenditure to control for subsidies. The remaining ones (iii, vi and viii) show results using state dummies instead.

Table 3.I.a: Distance as Explanatory Variable Using Human Capital and Share of Manufacturing in Regions' GDP – Equation 3.1

Dependent Variable Wi,t / Wsp,t	OLS			RE					FE	
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Dist. SP	-0.159	-0.16	-0.081	-0.185	-0.183	-0.119	-0.194	-0.144	-0.008	-0.317
	(0.007)***	(0.022)***	(0.018)***	(0.016)***	(0.020)***	(0.033)***	(0.021)***	(0.037)***	(0.025)	(0.052)***
SP Cont. Shock		0.008	-0.024		-0.008	-0.016	-0.008	-0.017		0.026
		(0.025)	(0.019)		(0.013)	(0.013)	(0.014)	(0.014)		(0.011)**
SP Exp. Shock		-0.002	-0.006		-0.001	-0.003	-0.002	-0.004		0.024
		(0.023)	(0.017)		(0.012)	(0.012)	(0.013)	(0.013)		(0.010)**
Dist. Port	-0.016	0.011	-0.003	0.001	0.04	0.029	0.04	0.03	0.023	0.048
	(0.005)***	(0.018)	(0.016)	(0.009)	(0.013)***	(0.014)**	(0.013)***	(0.014)**	(0.013)*	(0.015)***
Port Cont. Shock		-0.03	-0.035		-0.036	-0.036	-0.036	-0.036		-0.011
		(0.02)	(0.017)**		(0.011)***	(0.011)***	(0.011)***	(0.011)***		-0.009
Port Exp. Shock		-0.03	-0.043		-0.045	-0.045	-0.045	-0.045		-0.03
		(0.019)	(0.016)***		(0.010)***	(0.010)***	(0.010)***	(0.010)***		(0.008)***
Human Capital	0.093	0.093	0.072	0.093	0.093	0.08	0.096	0.081	0.027	0.018
	(0.006)***	(0.006)***	(0.006)***	(0.015)***	(0.015)***	(0.015)***	(0.014)***	(0.015)***	(0.03)	(0.033)
Subsidies	-0.061	-0.062		-0.018	-0.016		-0.016		-0.01	-0.011
	(0.006)***	(0.006)***		(0.006)***	(0.006)***		(0.006)***		(0.006)*	(0.006)*
Manuf. VA / GDP	0.134	0.134	0.144	0.092	0.093	0.096	0.092	0.096	0.085	0.078
	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***
Period Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Effect	No	No	Yes	No	No	Yes	No	Yes	No	No
Obs.	4.145	4.145	4.145	4.145	4.145	4.145	4.145	4.145	4.145	4.145
R-squared	0.47	0.47	0.54	0,38	0,38	0,51	0,39	0,51	0.07	0.09

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Tables 3.I.b shows results using market potential as explanatory variable but not instrumenting either human capital or market potential. The table has the same structure as Table 3.4, which is the following:

The first three columns (i to iii) presents OLS results, the following three random effects and the last two fixed effects. The first column of each method show results with not trade effect. Columns ii, v and viii use state government expenditure to control for subsidies. The remaining ones (iii and vii) show results using state dummies instead.

Table 3.I.b: Market Potential as Explanatory Variable Using Human Capital and Share of Manufacturing in Regions' GDP – Equation 3.5

Dependent Variable	OLS			RE			FE	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Dist. SP	-0.031	-0.021	-0.008	0.116	0.156	0.169	0.465	0.804
	(0.018)*	(0.053)	(0.049)	(0.046)**	(0.059)***	(0.067)**	(0.137)***	(0.168)***
SP Cont. Shock		-0.017	-0.017		-0.046	-0.043		-0.088
		(0.062)	(0.055)		(0.044)	(0.044)		(0.045)*
SP Exp. Shock		-0.007	-0.046		-0.038	-0.042		-0.029
		(0.058)	(0.052)		(0.042)	(0.042)		(0.042)
Dist. Port	0.245	0.21	0.153	0.14	0.066	0.033	-0.041	-0.11
	(0.015)***	(0.045)***	(0.044)***	(0.037)***	(0.048)	(0.050)	(0.060)	(0.066)*
Port Cont. Shock		0.027	0.08		0.082	0.09		0.117
		(0.053)	(0.048)*		(0.037)**	(0.037)**		(0.038)***
Port Exp. Shock		0.044	0.088		0.082	0.085		0.09
		(0.053)	(0.045)*		(0.036)**	(0.036)**		(0.036)**
Human Capital	0.099	0.099	0.08	0.093	0.094	0.077	0.03	0.044
	(0.006)***	(0.006)***	(0.006)***	(0.015)***	(0.014)***	(0.015)***	(0.034)	(0.035)
Subsidies	-0.063	-0.064		-0.016	-0.016		-0.009	-0.009
	(0.006)***	(0.006)***		(0.006)***	(0.006)***		(0.006)	(0.007)
Manuf. VA / GDP	0.137	0.137	0.143	0.094	0.094	0.095	0.083	0.084
	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***
Period Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Effect	No	No	Yes	No	No	Yes	No	Yes
Obs.	4.145	4.145	4.145	4.145	4.145	4.145	4.145	4.145
R-squared	0.47	0.47	0.55	0,38	0,38	0,51	0.1	0.1

* significant at 10%; ** significant at 5%; *** significant at 1%

Tables 3.I.c shows results using market potential as explanatory variable omitting human capital and instrumenting market potential. The table has similar structure as Table 3.4, which is the following:

The first three columns (i to iii) present 2SLS results, the following three IV random effects and the last two IV fixed effects. The first column of each method show results with no trade effect. Columns ii, v and viii use state government expenditure to control for subsidies. The remaining ones (iii and vii) show results using state dummies instead.

Table 3.I.c: Market Potential as Explanatory Variable Using Share of Manufacturing in Regions' GDP but not Human Capital – Equation 3.5

Panel A: Estimation Results								
Dependent Variable	2SLS			IVRE			IVFE	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Dist. SP	-0.751	-0.433	-0.146	-0.274	0.997	0.168	-0.426	-0.468
	(0.178)***	(0.603)	(0.133)	(0.157)*	(0.410)**	(0.054)***	-0.818	(0.185)**
SP Cont. Shock		0.573	0.164		-1.135	-0.238		0.289
		(0.780)	(0.144)		(0.590)*	(0.083)***		(0.109)***
SP Exp. Shock		0.439	-0.168		-1.007	-0.192		0.326
		(0.596)	(0.144)		(0.439)**	(0.063)***		(0.130)**
Dist. Port	1.04	0.409	0.167	0.264	-0.931	-0.184	-0.224	0.074
	(0.143)***	(0.591)	(0.146)	(0.122)**	(0.425)**	(0.060)***	(0.129)*	(0.043)*
Port Cont. Shock		-0.538	-0.190		1.536	0.299		-0.285
		(0.766)	(0.165)		(0.621)**	(0.091)***		(0.108)***
Port Exp. Shock		-0.411	-0.184		1.274	0.235		-0.322
		(0.586)	(0.156)		(0.460)***	(0.069)***		(0.128)**
Subsidies	-0.037	-0.127		0.036	0.033		-0.018	-0.081
	(0.008)***	(0.16)		-0.025	(0.013)**		(0.008)**	(0.027)***
Manuf. VA / GDP	0.135	0.118	0.146	0.104	0.096	0.113	0.084	0.082
	(0.008)***	(0.067)*	(0.027)***	(0.012)***	(0.007)***	(0.010)***	(0.006)***	(0.012)***
Period Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Effect	No	No	Yes	No	No	Yes	No	Yes
Obs.	4.145	4.145	4.145	4.145	4.145	4.145	4.145	4.145
R-squared	0,23	0,23	0,52	0,38	0,38	0,51	0,39	0,51
Panel B: First Stage Results								
Instruments								
Bank's Disburs.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Share	No	Yes	Yes	No	Yes	Yes	No	No
Distance	No	Yes	Yes	No	Yes	Yes	Yes	Yes
R-sq Int. MP	0.32	0.77	0.50	0.34	0.78	0.79	0.34	0.13
R-sq Ext. MP	0.35	0.88	0.82	0.39	0.88	0.88	0.30	0.07
Panel C: Over-identification Test								
P-value (Chi-sq)	0.94	0.50	0.19	0.49	0.60	0.97	0.37	0.67

Robust standard errors in parenthesis; * significant at 10%; ** significant at 5%; *** significant at 1%

Appendix 3.II: Further Results of Chapter 3
Outcomes using other Measures of Human Capital and Manufacturing Relevance

Tables 3.II.a until 3.II.h show results using distance as explanatory variable and it has the same structure as Table 3.3, which is the following:

The first six columns (i to vi) presents results with time-invariant distance to industry centre, while the last four columns (vii to x) shows using the distance to the industry centre weighted by its importance. The first column of each method (i, iv and ix) presents results with no trade effect. Columns ii, v, vii and x use state government expenditure to control for subsidies. The remaining ones (iii, vi and viii) show results using state dummies instead.

Table 3.II.a: Distance as Explanatory Variable Using Education and Share of Manufacturing in Regions' GDP – Equation 3.1

Dependent Variable	OLS			Random Effects					Fixed Effects	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Education	0.47	0.44	0.42	0.83	0.53	0.62	0.44	0.59	-0.29	-0.22
	(0.027)***	(0.025)***	(0.031)***	(0.111)***	(0.070)***	(0.105)***	(0.071)***	(0.104)***	(0.116)**	(0.23)
Distance to Sao Paulo	-0.11	-0.11	-0.10	-0.08	-0.11	-0.12	-0.15	-0.21	-0.18	-0.52
	(0.007)***	(0.014)***	(0.016)***	(0.024)***	(0.022)***	(0.031)***	(0.024)***	(0.032)***	(0.038)***	(0.067)***
Distance to Port	-0.03	0.00	0.00	0.04	0.04	0.04	0.03	0.04	0.02	0.05
	(0.005)***	(0.02)	(0.01)	(0.011)***	(0.013)***	(0.014)***	(0.012)**	(0.013)***	(0.01)	(0.016)***
Subsidies	-0.04	-0.05		0.01	-0.02		-0.02		-0.01	-0.01
	(0.006)***	(0.006)***		(0.01)	(0.006)***		(0.006)***		(0.01)	(0.006)*
Manuf. VA over GDP	0.12	0.13	0.14	0.06	0.09	0.10	0.09	0.09	0.08	0.08
	(0.005)***	(0.005)***	(0.005)***	(0.007)***	(0.006)***	(0.006)***	(0.006)***	(0.006)***	(0.007)***	(0.006)***
Sao Paulo 1st Shock		-0.01	-0.01		0.00	-0.01	-0.01	-0.01		0.02
		(0.01)	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)		(0.014)*
Sao Paulo 2nd Shock		-0.01	-0.01		0.00	-0.01	-0.01	-0.02		0.02
		(0.01)	(0.01)		(0.01)	(0.01)	(0.01)	(0.012)*		(0.02)
Port 1st Shock		-0.03	-0.03		-0.03	-0.03	-0.02	-0.02		-0.02
		(0.017)*	(0.02)		(0.010)***	(0.010)***	(0.009)**	(0.009)**		(0.009)**
Port 2nd Shock		-0.03	-0.04		-0.04	-0.04	-0.03	-0.03		-0.04
		(0.015)**	(0.015)**		(0.010)***	(0.010)***	(0.009)***	(0.009)***		(0.008)***
Constant	0.10	0.06	0.24	-0.09	0.04	0.16	-0.04	0.36	-0.43	1.04
	(0.053)*	-0.05	(0.092)***	(0.14)	(0.12)	(0.24)	(0.09)	(0.185)*	(0.166)***	(0.238)***
Observations	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145
R-squared	0.47	0.49	0.55	0.38	0.47	0.52	0.47	0.53	0.23	0.31

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.II.b: Distance as Explanatory Variable Using Productivity and Share of Manufacturing in Regions' GDP – Equation 3.1

Dependent Variable	OLS			Random Effects					Fixed Effects	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Productivity	1.11	1.32	1.84	0.42	0.80	0.86	0.59	0.79	0.12	0.13
	(0.103)***	(0.144)***	(0.372)***	(0.070)***	(0.243)***	(0.298)***	(0.221)***	(0.286)***	(0.046)***	(0.13)
Distance to Sao Paulo	-0.09	-0.15	-0.13	-0.19	-0.25	-0.20	-0.32	-0.38	-0.10	-0.48
	(0.013)***	(0.025)***	(0.043)***	(0.021)***	(0.033)***	(0.056)***	(0.027)***	(0.046)***	(0.020)***	(0.046)***
Distance to Port	-0.02	-0.06	-0.02	-0.01	-0.02	-0.02	0.00	-0.01	0.03	0.05
	(0.008)***	(0.028)**	(0.04)	(0.01)	(0.02)	(0.03)	(0.02)	(0.02)	(0.012)**	(0.016)***
Subsidies	-0.01	0.02		-0.03	-0.03		-0.03		-0.01	-0.02
	(0.01)	(0.01)		(0.007)***	(0.009)***		(0.007)***		(0.006)*	(0.007)**
Manuf. VA over GDP	-0.31	-0.41	-0.65	-0.10	-0.34	-0.37	-0.23	-0.34	0.02	0.01
	(0.043)***	(0.062)***	(0.163)***	(0.032)***	(0.132)***	(0.161)**	(0.120)*	(0.155)**	(0.02)	(0.07)
Sao Paulo 1st Shock		0.06	0.09		0.05	0.04	0.05	0.05		0.02
		(0.027)**	(0.040)**		(0.015)***	(0.017)**	(0.017)***	(0.021)***		(0.013)*
Sao Paulo 2nd Shock		0.08	0.09		0.04	0.05	0.04	0.05		0.02
		(0.026)***	(0.037)**		(0.012)***	(0.014)***	(0.014)***	(0.018)***		(0.01)
Port 1st Shock		0.07	0.12		0.04	0.05	0.03	0.06		-0.01
		(0.033)**	(0.051)**		(0.024)*	(0.030)*	(0.02)	(0.029)**		(0.02)
Port 2nd Shock		0.04	0.10		0.04	0.05	0.02	0.04		-0.03
		(0.03)	(0.048)**		(0.03)	(0.03)	(0.02)	(0.03)		(0.015)*
Constant	0.21	0.33	-0.44	0.37	0.40	0.03	0.25	0.59	-0.66	0.97
	(0.084)**	(0.098)***	(0.29)	(0.144)***	(0.216)*	(0.46)	(0.126)*	(0.312)*	(0.112)***	(0.211)***
Observations	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145
R-squared	0.27	0.44	0.70	0.62	0.46	0.46	0.53	0.49	0.59	0.43

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.II.c: Distance as Explanatory Variable Using Human Capital and Share of Manufacturing Employment – Equation 3.1

Dependent Variable	OLS			Random Effects					Fixed Effects	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Human Capital	0.46	0.44	0.25	0.48	0.40	0.26	0.30	0.23	0.34	0.12
	(0.026)***	(0.024)***	(0.013)***	(0.067)***	(0.052)***	(0.032)***	(0.049)***	(0.031)***	(0.164)**	(0.10)
Distance to Sao Paulo	-0.12	-0.12	-0.07	-0.14	-0.16	-0.11	-0.27	-0.33	-0.13	-0.65
	(0.010)***	(0.019)***	(0.019)***	(0.028)***	(0.028)***	(0.039)***	(0.027)***	(0.035)***	(0.052)**	(0.056)***
Distance to Port	0.06	0.08	0.02	0.06	0.07	0.04	0.05	0.04		0.06
	(0.009)***	(0.021)***	(0.02)	(0.013)***	(0.015)***	(0.015)***	(0.014)***	(0.014)***		(0.016)***
Subsidies	-0.04	-0.05		-0.01	-0.02		-0.02		-0.01	-0.02
	(0.008)***	(0.007)***		(0.006)**	(0.006)***		(0.006)***		(0.01)	(0.006)***
Employment Share	0.03	0.04	0.04	-0.06	-0.03	-0.03	-0.05	-0.04	-0.09	-0.08
	(0.007)***	(0.007)***	(0.006)***	(0.009)***	(0.008)***	(0.008)***	(0.008)***	(0.008)***	(0.011)***	(0.009)***
Sao Paulo 1st Shock		0.00	0.01		0.01	0.01	0.01	0.01		0.01
		(0.02)	(0.02)		(0.01)	(0.01)	(0.01)	(0.01)		(0.01)
Sao Paulo 2nd Shock		0.00	0.01		0.01	0.01	0.00	0.00		0.00
		(0.02)	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)		(0.01)
Port 1st Shock		-0.03	-0.03		-0.02	-0.02	-0.01	-0.01		0.00
		(0.02)	(0.02)		(0.011)**	(0.010)*	(0.01)	(0.01)		(0.01)
Port 2nd Shock		-0.03	-0.04		-0.03	-0.03	-0.02	-0.02		-0.02
		(0.02)	(0.017)**		(0.010)***	(0.010)***	(0.009)**	(0.009)**		(0.009)**
Constant	1.43	1.34	0.62	1.75	1.52	0.63	1.26	1.47	0.99	2.10
	(0.087)***	(0.084)***	(0.108)***	(0.211)***	(0.183)***	(0.295)**	(0.174)***	(0.232)***	(0.562)*	(0.419)***
Observations	4,146	4,146	4,146	4,146	4,146	4,146	4,146	4,146	4,146	4,146
R-squared	0.01	0.06	0.39	0.22	0.26	0.37	0.32	0.37	0.22	0.30

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.II.d: Distance as Explanatory Variable Using Education and Share of Manufacturing Employment – Equation 3.1

Dependent Variable	OLS			Random Effects					Fixed Effects	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Education	0.66	0.61	0.68	1.19	0.80	1.10	0.66	1.03	-0.20	-0.30
	(0.029)***	(0.027)***	(0.033)***	(0.139)***	(0.083)***	(0.130)***	(0.082)***	(0.128)***	(0.13)	(0.24)
Distance to Sao Paulo	-0.14	-0.13	-0.12	-0.10	-0.12	-0.14	-0.19	-0.27	-0.27	-0.74
	(0.008)***	(0.015)***	(0.018)***	(0.027)***	(0.025)***	(0.036)***	(0.027)***	(0.036)***	(0.035)***	(0.065)***
Distance to Port	-0.04	-0.01	-0.01	0.04	0.03	0.05	0.02	0.05	0.04	0.05
	(0.005)***	(0.02)	(0.02)	(0.012)***	(0.014)**	(0.016)***	(0.013)*	(0.014)***	(0.015)**	(0.016)***
Subsidies	-0.03	-0.04		0.00	-0.02		-0.02		-0.01	-0.02
	(0.006)***	(0.006)***		(0.01)	(0.006)***		(0.006)***		(0.006)**	(0.006)**
Employment Share	0.00	0.01	0.04	-0.13	-0.04	-0.04	-0.05	-0.06	-0.06	-0.07
	(0.01)	(0.006)**	(0.006)***	(0.014)***	(0.008)***	(0.009)***	(0.008)***	(0.009)***	(0.012)***	(0.010)***
Sao Paulo 1st Shock		0.00	-0.01		-0.01	-0.02	-0.01	-0.03		0.03
		(0.02)	(0.02)		(0.01)	(0.010)*	(0.01)	(0.013)**		(0.014)**
Sao Paulo 2nd Shock		-0.01	-0.01		-0.01	-0.02	-0.03	-0.05		0.02
		(0.01)	(0.01)		(0.01)	(0.010)**	(0.012)**	(0.013)***		(0.02)
Port 1st Shock		-0.03	-0.03		-0.02	-0.02	-0.01	-0.01		0.00
		(0.018)*	(0.017)*		(0.011)*	(0.011)*	(0.01)	(0.01)		(0.01)
Port 2nd Shock		-0.03	-0.04		-0.02	-0.02	-0.02	-0.01		-0.02
		(0.017)*	(0.016)**		(0.010)**	(0.010)**	(0.01)	(0.01)		(0.009)***
Constant	0.34	0.28	0.41	-0.02	0.10	0.27	0.10	0.66	-0.22	1.72
	(0.058)***	(0.057)***	(0.102)***	(0.16)	(0.14)	(0.28)	(0.10)	(0.211)***	(0.15)	(0.241)***
Observations	4,146	4,146	4,146	4,146	4,146	4,146	4,146	4,146	4,146	4,146
R-squared	0.37	0.39	0.45	0.26	0.35	0.39	0.36	0.39	0.11	0.20

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.II.e: Distance as Explanatory Variable Using Productivity and Share of Manufacturing Employment – Equation 3.1

Dependent Variable	OLS			Random Effects					Fixed Effects	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Productivity	0.63	0.67	0.62	0.47	0.53	0.48	0.50	0.47	0.09	0.12
	(0.026)***	(0.029)***	(0.029)***	(0.035)***	(0.044)***	(0.039)***	(0.048)***	(0.041)***	(0.055)*	(0.09)
Distance to Sao Paulo	-0.08	-0.10	-0.09	-0.11	-0.13	-0.11	-0.15	-0.17	-0.19	-0.59
	(0.009)***	(0.015)***	(0.017)***	(0.017)***	(0.020)***	(0.026)***	(0.023)***	(0.030)***	(0.036)***	(0.084)***
Distance to Port	-0.01	-0.02	0.00	-0.02	-0.01	0.00	-0.01	0.00	0.04	0.04
	(0.005)***	(0.02)	(0.02)	(0.008)*	(0.01)	(0.01)	(0.01)	(0.01)	(0.014)***	(0.017)***
Subsidies	-0.03	-0.02		-0.03	-0.02		-0.02		-0.01	-0.02
	(0.006)***	(0.006)***		(0.006)***	(0.006)***		(0.006)***		(0.006)**	(0.006)***
Employment Share	0.00	-0.01	-0.01	0.03	0.01	0.01	0.00	0.00	-0.05	-0.06
	(0.01)	(0.006)*	(0.01)	(0.007)***	(0.01)	(0.01)	(0.01)	(0.01)	(0.015)***	(0.015)***
Sao Paulo 1st Shock		0.03	0.02		0.02	0.02	0.03	0.02		0.02
		(0.015)*	(0.01)		(0.010)**	(0.009)**	(0.012)**	(0.012)**		(0.010)*
Sao Paulo 2nd Shock		0.04	0.03		0.03	0.03	0.03	0.03		0.01
		(0.014)***	(0.013)***		(0.009)***	(0.009)***	(0.012)***	(0.011)***		(0.01)
Port 1st Shock		0.02	0.02		0.01	0.01	0.02	0.02		0.00
		(0.02)	(0.02)		(0.01)	(0.01)	(0.01)	(0.010)*		(0.01)
Port 2nd Shock		0.00	0.01		0.00	0.00	0.00	0.00		-0.02
		(0.02)	(0.02)		(0.01)	(0.01)	(0.01)	(0.01)		(0.009)*
Constant	-0.09	-0.07	-0.16	0.05	0.04	-0.08	-0.09	0.01	-0.42	1.35
	(0.06)	(0.06)	(0.10)	(0.10)	(0.12)	(0.200	(0.09)	(0.17)	(0.125)***	(0.311)***
Observations	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145
R-squared	0.46	0.42	0.52	0.65	0.65	0.68	0.65	0.68	0.42	0.38

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.II.f: Distance as Explanatory Variable Using Human Capital and Percentage of Number of Manufacturing Firms – Equation 3.1

Dependent Variable	OLS			Random Effects					Fixed Effects	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Human Capital	-3.86	-2.00	1.04	0.39	0.28	0.18	0.17	0.21	0.40	0.17
	(1.277)***	(0.377)***	(0.103)***	(0.128)***	(0.091)***	(0.072)**	(0.087)*	(0.066)***	(0.159)**	(0.104)*
Distance to Sao Paulo	0.19	0.02	-0.12	-0.20	-0.21	-0.16	-0.40	-0.47	-0.09	-0.55
	(0.14)	(0.07)	(0.030)***	(0.047)***	(0.046)***	(0.082)**	(0.038)***	(0.044)***	(0.041)**	(0.054)***
Distance to Port	-0.43	-0.06	0.01	0.04	0.06	0.06	0.06	0.06		0.06
	(0.135)***	(0.06)	(0.03)	(0.013)***	(0.015)***	(0.015)***	(0.014)***	(0.014)***		(0.016)***
Subsidies	-0.08	-0.10		-0.01	-0.02		-0.02		-0.01	-0.02
	(0.039)*	(0.022)***		(0.006)**	(0.006)***		(0.006)**		(0.01)	(0.006)**
Number of Firms Share	2.25	1.25	-0.55	-0.08	-0.02	-0.01	-0.03	-0.03	-0.09	-0.03
	(0.719)***	(0.221)***	(0.063)***	(0.025)***	(0.02)	(0.02)	(0.016)*	(0.015)**	(0.021)***	(0.014)**
Sao Paulo 1st Shock		0.04	0.00		0.01	0.01	0.02	0.01		0.02
		(0.05)	(0.03)		(0.01)	(0.01)	(0.01)	(0.01)		(0.011)*
Sao Paulo 2nd Shock		0.02	0.02		0.02	0.02	0.01	0.01		0.00
		(0.05)	(0.02)		(0.008)*	(0.008)**	(0.01)	(0.01)		(0.01)
Port 1st Shock		-0.18	0.04		-0.02	-0.02	-0.01	-0.01		-0.01
		(0.068)***	(0.03)		(0.010)**	(0.010)**	(0.01)	(0.01)		(0.01)
Port 2nd Shock		-0.19	0.03		-0.03	-0.03	-0.03	-0.03		-0.03
		(0.066)***	(0.03)		(0.009)***	(0.009)***	(0.008)***	(0.008)***		(0.009)***
Constant	-6.40	-3.17	1.47	1.44	1.19	0.48	1.11	1.85	0.69	1.77
	(2.347)***	(0.738)***	(0.198)***	(0.390)***	(0.335)***	(0.65)	(0.306)***	(0.407)***	(0.51)	(0.425)***
Observations	4,146	4,146	4,146	4,146	4,146	4,146	4,146	4,146	4,146	4,146
R-squared	0.71	0.86	0.35	0.28	0.31	0.40	0.34	0.38	0.21	0.33

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.II.g: Distance as Explanatory Variable Using Education and Percentage of Number of Manufacturing Firms – Equation 3.1

Dependent Variable	OLS			Random Effects					Fixed Effects	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Education	0.70	0.62	0.76	1.75	0.77	1.23	0.65	1.20	-0.26	-0.38
	(0.037)***	(0.034)***	(0.045)***	(0.237)***	(0.097)***	(0.185)***	(0.096)***	(0.178)***	(0.128)**	(0.229)*
Distance to Sao Paulo	-0.14	-0.13	-0.13	-0.11	-0.12	-0.15	-0.18	-0.24	-0.22	-0.66
	(0.008)***	(0.015)***	(0.018)***	(0.029)***	(0.025)***	(0.036)***	(0.027)***	(0.037)***	(0.031)***	(0.065)***
Distance to Port	-0.04	-0.01	-0.01	0.05	0.04	0.06	0.03	0.05	0.03	0.06
	(0.005)***	(0.02)	(0.02)	(0.013)***	(0.014)***	(0.016)***	(0.013)**	(0.015)***	(0.015)**	(0.017)***
Subsidies	-0.02	-0.04		0.01	-0.02		-0.02		-0.01	-0.01
	(0.006)***	(0.006)***		(0.01)	(0.006)**		(0.006)***		(0.006)*	(0.006)**
Number of Firms Share	-0.02	0.01	-0.01	-0.23	-0.01	-0.04	-0.01	-0.05	-0.01	-0.01
	(0.007)**	(0.01)	(0.01)	(0.034)***	(0.01)	(0.016)**	(0.01)	(0.015)***	(0.02)	(0.01)
Sao Paulo 1st Shock		0.00	-0.01		0.00	-0.02	-0.01	-0.03		0.04
		(0.02)	(0.02)		(0.01)	(0.01)	(0.01)	(0.014)**		(0.014)***
Sao Paulo 2nd Shock		-0.01	-0.01		-0.01	-0.02	-0.02	-0.06		0.03
		(0.01)	(0.01)		(0.01)	(0.011)**	(0.012)**	(0.015)***		(0.016)*
Port 1st Shock		-0.03	-0.02		-0.03	-0.02	-0.02	-0.01		-0.01
		(0.02)	(0.02)		(0.011)**	(0.011)**	(0.010)*	(0.01)		(0.01)
Port 2nd Shock		-0.03	-0.03		-0.03	-0.03	-0.02	-0.02		-0.04
		(0.017)*	(0.016)*		(0.010)***	(0.011)**	(0.009)**	(0.010)*		(0.009)***
Constant	0.31	0.31	0.43	-0.73	0.05	0.23	0.00	0.42	-0.47	1.31
	(0.058)***	(0.057)***	(0.102)***	(0.238)***	(0.14)	(0.28)	(0.11)	(0.217)*	(0.177)***	(0.247)***
Observations	4,146	4,146	4,146	4,146	4,146	4,146	4,146	4,146	4,146	4,146
R-squared	0.36	0.39	0.44	0.20	0.37	0.39	0.38	0.40	0.07	0.19

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.II.h: Distance as Explanatory Variable Using Productivity and Percentage of Number of Manufacturing Firms – Equation 3.1

Dependent Variable	OLS			Random Effects					Fixed Effects	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Productivity	0.65	0.75	0.70	0.39	0.49	0.45	0.46	0.44	0.12	0.15
	(0.033)***	(0.045)***	(0.042)***	(0.034)***	(0.057)***	(0.049)***	(0.061)***	(0.050)***	(0.052)**	(0.080)*
Distance to Sao Paulo	-0.08	-0.11	-0.09	-0.11	-0.13	-0.11	-0.17	-0.18	-0.12	-0.49
	(0.009)***	(0.017)***	(0.018)***	(0.018)***	(0.022)***	(0.028)***	(0.025)***	(0.031)***	(0.034)***	(0.068)***
Distance to Port	-0.01	-0.03	0.00	-0.01	-0.01	0.00	-0.01	0.00	0.03	0.05
	(0.005)***	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.013)**	(0.017)***
Subsidies	-0.03	-0.02		-0.02	-0.02		-0.02		-0.01	-0.02
	(0.006)***	(0.007)***		(0.006)***	(0.006)***		(0.006)***		(0.006)**	(0.006)***
Number of Firms Share	-0.01	-0.04	-0.03	0.05	0.01	0.01	0.00	0.01	-0.01	-0.01
	(0.01)	(0.009)***	(0.008)***	(0.007)***	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Sao Paulo 1st Shock		0.03	0.02		0.02	0.02	0.03	0.02		0.03
		(0.017)*	(0.02)		(0.009)**	(0.009)**	(0.012)**	(0.011)**		(0.010)**
Sao Paulo 2nd Shock		0.04	0.04		0.03	0.03	0.03	0.03		0.02
		(0.016)***	(0.014)***		(0.009)***	(0.009)***	(0.011)***	(0.011)***		(0.010)*
Port 1st Shock		0.03	0.03		0.01	0.01	0.01	0.02		0.00
		(0.02)	(0.018)*		(0.01)	(0.01)	(0.01)	(0.01)		(0.01)
Port 2nd Shock		0.01	0.02		0.00	-0.01	0.00	0.00		-0.02
		(0.02)	(0.02)		(0.01)	(0.01)	(0.01)	(0.01)		(0.009)***
Constant	-0.11	-0.18	-0.23	0.21	0.08	-0.05	-0.05	0.08	-0.67	0.94
	(0.065)*	(0.076)**	(0.113)**	(0.106)**	(0.13)	(0.21)	(0.10)	(0.18)	(0.115)***	(0.247)***
Observations	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145
R-squared	0.42	0.26	0.41	0.67	0.65	0.69	0.66	0.69	0.55	0.43

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Annex 3.III: Further Results using other Measures of Human Capital and Manufacturing Relevance of Chapter 3

Tables 3.III.a until 3.III.h show results using market potential as explanatory variable and it has the same structure as Table 3.4, which is the following:

The first three columns show results using two-stage least squares, where the first with no trade shock considered, the second using government expenditures and third using state dummies. Random effects outcomes are presented from iv to vi, following the same structure from 2SLS, which means iv no trade effect, v government expenditure and vi state dummies. The last two columns show fixed effects results using government expenditure only, with and without trade shock.

Table 3.III.a: Market Potential as Explanatory Variable Using Education and Share of Manufacturing in Regions' GDP – Equation 3.5

Dependent Variable	2SLS			Random Effects			Fixed Effects	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Internal Mkt Potential	0.18	0.26	0.24	0.41	0.49	0.46	0.73	1.81
	(0.041)***	(0.136)*	(0.128)*	(0.100)***	(0.123)***	(0.129)***	(2.10)	(1.93)
External Mkt Potential	0.04	-0.06	-0.01	-0.19	-0.29	-0.18	-0.15	-0.26
	(0.03)	(0.13)	(0.12)	(0.083)**	(0.105)***	(0.097)*	(0.17)	(0.17)
Education	0.47	0.44	0.40	0.67	0.53	0.42	-0.13	0.09
	(0.027)***	(0.026)***	(0.033)***	(0.097)***	(0.083)***	(0.093)***	(0.12)	(0.32)
Subsidies	-0.04	-0.05		0.00	-0.02		-0.01	0.00
	(0.006)***	(0.006)***		(0.01)	(0.006)***		(0.01)	(0.01)
Manuf. VA over GDP	0.12	0.12	0.13	0.07	0.09	0.10	0.08	0.09
	(0.005)***	(0.005)***	(0.005)***	(0.007)***	(0.006)***	(0.006)***	(0.011)***	(0.010)***
Internal Mkt Potential 1st Shock		-0.14	-0.12		-0.19	-0.17		-0.21
		(0.16)	(0.15)		(0.094)**	(0.095)*		(0.096)**
Internal Mkt Potential 2nd Shock		-0.06	-0.22		-0.24	-0.26		-0.26
		(0.15)	(0.14)		(0.088)***	(0.087)***		(0.092)***
External Mkt Potential 1st Shock		0.16	0.15		0.21	0.19		0.23
		(0.16)	(0.15)		(0.092)**	(0.093)**		(0.094)**
External Mkt Potential 2nd Shock		0.09	0.25		0.26	0.29		0.27
		(0.15)	(0.140)*		(0.086)***	(0.085)***		(0.097)***
Constant	0.63	0.60	1.11	0.86	0.76	1.41	2.73	9.39
	(0.094)***	(0.093)***	(0.208)***	(0.248)***	(0.252)***	(0.530)***	(14.50)	(13.01)
Observations	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145
R-squared	0.45	0.47	0.54	0.37	0.44	0.52	0.26	0.28

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.III.b: Market Potential as Explanatory Variable Using Productivity and Share of Manufacturing in Regions' GDP – Equation 3.5

Dependent Variable	2SLS			Random Effects			Fixed Effects	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Internal Mkt Potential	0.01	0.07	0.28	0.28	-0.13	0.03	1.57	1.97
	(0.05)	(0.16)	(0.148)*	(0.100)***	(0.25)	(0.40)	(2.65)	(1.64)
External Mkt Potential	0.18	0.18	0.03	0.04	0.52	0.43	-0.20	-0.20
	(0.037)***	(0.16)	(0.14)	(0.08)	(0.238)**	(0.37)	(0.14)	(0.20)
Productivity	0.83	0.87	0.82	0.64	1.28	1.81	0.10	-0.11
	(0.055)***	(0.060)***	(0.084)***	(0.089)***	(0.268)***	(0.863)**	(0.09)	(0.42)
Subsidies	-0.02	-0.01		-0.04	-0.03		-0.01	0.00
	(0.007)***	(0.01)		(0.008)***	(0.011)**		(0.01)	(0.02)
Manuf. VA over GDP	-0.20	-0.22	-0.21	-0.19	-0.53	-0.82	0.04	0.16
	(0.023)***	(0.026)***	(0.037)***	(0.040)***	(0.134)***	(0.441)*	(0.04)	(0.23)
Internal Mkt Potential 1st Shock		0.10	0.15		0.39	0.72		-0.26
		(0.19)	(0.18)		(0.225)*	(0.50)		(0.20)
Internal Mkt Potential 2nd Shock		-0.18	-0.01		0.37	0.78		-0.33
		(0.18)	(0.17)		(0.23)	(0.57)		(0.32)
External Mkt Potential 1st Shock		-0.17	-0.22		-0.52	-0.89		0.28
		(0.19)	(0.17)		(0.239)**	(0.57)		(0.24)
External Mkt Potential 2nd Shock		0.10	-0.06		-0.49	-0.94		0.35
		(0.18)	(0.17)		(0.248)**	(0.64)		(0.36)
Constant	0.56	0.64	1.15	1.17	1.18	1.43	8.04	10.86
	(0.110)***	(0.109)***	(0.242)***	(0.218)***	(0.311)***	(1.17)	(17.98)	(10.05)
Observations	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145
R-squared	0.26	0.26	0.40	0.58	0.42	0.33	0.30	0.24

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.III.c: Market Potential as Explanatory Variable Using Human Capital and Share of Manufacturing Employment – Equation 3.5

Dependent Variable	2SLS			Random Effects			Fixed Effects	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Internal Mkt Potential	0.12	0.53	0.57	0.44	0.62	0.58	0.81	1.17
	(0.044)***	(0.150)***	(0.142)***	(0.103)***	(0.131)***	(0.160)***	(2.12)	(1.13)
External Mkt Potential	0.22	-0.22	-0.22	0.00	-0.22	-0.20	-0.31	-0.40
	(0.034)***	(0.15)	(0.14)	(0.08)	(0.110)**	(0.109)*	(0.155)**	(0.152)***
Human Capital	0.13	0.13	0.12	0.11	0.13	0.12	0.01	0.00
	(0.007)***	(0.007)***	(0.007)***	(0.018)***	(0.020)***	(0.022)***	(0.10)	(0.08)
Subsidies	-0.06	-0.06		-0.02	-0.02		-0.02	-0.02
	(0.006)***	(0.006)***		(0.006)***	(0.006)***		(0.013)*	(0.009)*
Employment Share	0.03	0.04	0.05	-0.04	-0.03	-0.03	-0.04	-0.04
	(0.006)***	(0.006)***	(0.007)***	(0.007)***	(0.008)***	(0.008)***	(0.010)***	(0.009)***
Internal Mkt Potential 1st Shock		-0.33	-0.31		-0.25	-0.18		-0.16
		(0.178)*	(0.168)*		(0.096)**	(0.096)*		(0.10)
Internal Mkt Potential 2nd Shock		-0.56	-0.53		-0.38	-0.29		-0.20
		(0.169)***	(0.161)***		(0.088)***	(0.086)***		(0.078)**
External Mkt Potential 1st Shock		0.34	0.32		0.25	0.18		0.16
		(0.174)**	(0.165)*		(0.094)***	(0.094)*		(0.096)*
External Mkt Potential 2nd Shock		0.56	0.53		0.38	0.29		0.19
		(0.166)***	(0.158)***		(0.086)***	(0.085)***		(0.076)**
Constant	1.59	1.50	2.02	2.27	2.14	2.14	2.17	3.88
	(0.096)***	(0.098)***	(0.228)***	(0.226)***	(0.249)***	(0.706)***	(14.14)	(7.21)
Observations	4,146	4,146	4,146	4,146	4,146	4,146	4,146	4,146
R-squared	0.35	0.35	0.43	0.29	0.31	0.40	0.15	0.19

Robust Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.III.d: Market Potential as Explanatory Variable Using Education and Share of Manufacturing Employment – Equation 3.5

Dependent Variable	2SLS			Random Effects			Fixed Effects	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Internal Mkt Potential	0.23	0.25	0.25	0.48	0.54	0.69	1.30	1.15
	(0.043)***	(0.149)*	(0.141)*	(0.106)***	(0.130)***	(0.142)***	(2.83)	(2.08)
External Mkt Potential	0.04	-0.01	0.03	-0.17	-0.30	-0.28	-0.28	-0.41
	(0.03)	(0.14)	(0.13)	(0.085)**	(0.110)***	(0.105)***	(0.19)	(0.176)**
Education	0.68	0.63	0.65	0.82	0.83	0.71	-0.02	0.00
	(0.030)***	(0.028)***	(0.035)***	(0.101)***	(0.091)***	(0.106)***	(0.17)	(0.36)
Subsidies	-0.03	-0.04		-0.01	-0.02		-0.02	-0.02
	(0.006)***	(0.006)***		(0.01)	(0.006)***		(0.02)	(0.01)
Employment Share	-0.01	0.00	0.03	-0.10	-0.05	-0.04	-0.04	-0.04
	(0.006)*	(0.01)	(0.007)***	(0.011)***	(0.008)***	(0.008)***	(0.015)**	(0.009)***
Internal Mkt Potential 1st Shock		-0.09	-0.14		-0.13	-0.13		-0.16
		(0.18)	(0.17)		(0.10)	(0.10)		(0.10)
Internal Mkt Potential 2nd Shock		0.03	-0.17		-0.11	-0.16		-0.20
		(0.17)	(0.16)		(0.09)	(0.091)*		(0.097)**
External Mkt Potential 1st Shock		0.12	0.17		0.15	0.15		0.16
		(0.17)	(0.16)		(0.10)	(0.10)		(0.096)*
External Mkt Potential 2nd Shock		0.01	0.21		0.14	0.19		0.20
		(0.17)	(0.15)		(0.09)	(0.089)**		(0.104)*
Constant	0.99	0.96	1.51	1.27	1.01	2.21	5.50	3.75
	(0.102)***	(0.101)***	(0.230)***	(0.267)***	(0.273)***	(0.604)***	(19.51)	(13.96)
Observations	4,146	4,146	4,146	4,146	4,146	4,146	4,146	4,146
R-squared	0.34	0.36	0.45	0.28	0.31	0.38	0.19	0.18

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.III.e: Market Potential as Explanatory Variable Using Productivity and Share of Manufacturing Employment – Equation 3.5

Dependent Variable	2SLS			Random Effects			Fixed Effects	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Internal Mkt Potential	0.04	0.14	0.24	0.15	0.15	0.19	1.24	1.16
	(0.04)	(0.13)	(0.116)**	(0.076)*	(0.12)	(0.12)	(2.42)	(1.30)
External Mkt Potential	0.12	0.04	-0.02	0.05	0.06	0.04	-0.28	-0.41
	(0.027)***	(0.12)	(0.11)	(0.06)	(0.10)	(0.09)	(0.172)*	(0.182)**
Productivity	0.56	0.57	0.51	0.48	0.50	0.44	0.02	0.00
	(0.019)***	(0.020)***	(0.022)***	(0.032)***	(0.035)***	(0.038)***	(0.12)	(0.15)
Subsidies	-0.04	-0.03		-0.03	-0.02		-0.02	-0.02
	(0.005)***	(0.005)***		(0.006)***	(0.006)***		(0.02)	(0.01)
Employment Share	0.01	0.00	0.00	0.02	0.00	0.00	-0.04	-0.04
	(0.01)	(0.01)	(0.01)	(0.007)***	(0.01)	(0.01)	(0.018)**	(0.015)***
Internal Mkt Potential 1st Shock		-0.02	0.04		0.01	0.03		-0.16
		(0.15)	(0.14)		(0.11)	(0.10)		(0.11)
Internal Mkt Potential 2nd Shock		-0.20	-0.13		-0.12	-0.11		-0.20
		(0.14)	(0.13)		(0.10)	(0.10)		(0.118)*
External Mkt Potential 1st Shock		-0.02	-0.08		-0.04	-0.06		0.16
		(0.14)	(0.13)		(0.11)	(0.10)		(0.11)
External Mkt Potential 2nd Shock		0.16	0.09		0.09	0.08		0.20
		(0.14)	(0.13)		(0.10)	(0.09)		(0.12)
Constant	0.30	0.32	0.78	0.54	0.52	0.76	5.14	3.82
	(0.093)***	(0.093)***	(0.202)***	(0.188)***	(0.192)***	(0.435)*	(16.76)	(8.87)
Observations	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145
R-squared	0.55	0.55	0.63	0.62	0.65	0.69	0.21	0.18

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.III.f: Market Potential as Explanatory Variable Using Human Capital and Percentage of Number of Manufacturing Firms – Equation 3.5

Dependent Variable	2SLS			Random Effects			Fixed Effects	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Internal Mkt Potential	0.09	0.37	0.53	0.47	0.64	0.59	-0.12	2.04
	(0.046)**	(0.151)**	(0.144)***	(0.105)***	(0.133)***	(0.168)***	(1.76)	(1.187)*
External Mkt Potential	0.26	-0.04	-0.14	-0.03	-0.24	-0.22	-0.39	-0.42
	(0.034)***	(0.15)	(0.14)	(0.08)	(0.110)**	(0.111)**	(0.151)***	(0.155)***
Human Capital	0.14	0.14	0.14	0.15	0.14	0.13	0.14	0.05
	(0.011)***	(0.011)***	(0.012)***	(0.023)***	(0.024)***	(0.028)***	(0.11)	(0.08)
Subsidies	-0.06	-0.06		-0.02	-0.02		-0.02	-0.01
	(0.006)***	(0.006)***		(0.006)***	(0.006)***		(0.012)*	(0.01)
Number of Firms Share	-0.01	0.00	-0.02	-0.03	-0.01	-0.01	-0.03	0.00
	(0.01)	(0.01)	(0.010)*	(0.010)***	(0.01)	(0.01)	(0.018)*	(0.01)
Internal Mkt Potential 1st Shock		-0.15	-0.20		-0.24	-0.19		-0.22
		(0.18)	(0.17)		(0.095)**	(0.094)**		(0.098)**
Internal Mkt Potential 2nd Shock		-0.44	-0.44		-0.38	-0.30		-0.23
		(0.170)***	(0.162)***		(0.087)***	(0.084)***		(0.078)***
External Mkt Potential 1st Shock		0.16	0.21		0.25	0.20		0.22
		(0.17)	(0.17)		(0.093)***	(0.092)**		(0.095)**
External Mkt Potential 2nd Shock		0.45	0.44		0.38	0.30		0.23
		(0.166)***	(0.158)***		(0.085)***	(0.082)***		(0.076)***
Constant	1.68	1.58	2.31	2.31	2.11	2.06	-4.11	9.82
	(0.100)***	(0.102)***	(0.237)***	(0.230)***	(0.257)***	(0.744)***	(11.76)	(7.64)
Observations	4,146	4,146	4,146	4,146	4,146	4,146	4,146	4,146
R-squared	0.35	0.34	0.42	0.29	0.32	0.41	0.16	0.23

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.III.g: Market Potential as Explanatory Variable Using Education and Percentage of Number of Manufacturing Firms – Equation 3.5

Dependent Variable	2SLS			Random Effects			Fixed Effects	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Internal Mkt Potential	0.16	0.05	0.16	0.83	0.58	0.72	2.81	3.22
	(0.046)***	(0.15)	(0.14)	(0.141)***	(0.132)***	(0.147)***	(2.06)	(2.13)
External Mkt Potential	0.10	0.17	0.12	-0.57	-0.36	-0.32	-0.20	-0.35
	(0.036)***	(0.14)	(0.13)	(0.129)***	(0.112)***	(0.108)***	(0.17)	(0.179)**
Education	0.72	0.66	0.73	2.06	0.91	0.87	-0.18	-0.21
	(0.038)***	(0.035)***	(0.043)***	(0.288)***	(0.111)***	(0.137)***	(0.15)	(0.34)
Subsidies	-0.03	-0.04		0.00	-0.02		-0.01	0.00
	(0.006)***	(0.006)***		(0.01)	(0.006)***		(0.01)	(0.02)
Number of Firms Share	-0.02	0.00	-0.01	-0.28	-0.03	-0.03	0.01	0.00
	(0.008)***	(0.01)	(0.01)	(0.041)***	(0.013)***	(0.014)**	(0.02)	(0.01)
Internal Mkt Potential 1st Shock		0.05	-0.06		-0.14	-0.14		-0.23
		(0.17)	(0.17)		(0.10)	(0.10)		(0.102)**
Internal Mkt Potential 2nd Shock		0.15	-0.10		-0.09	-0.15		-0.20
		(0.17)	(0.16)		(0.09)	(0.09)		(0.096)**
External Mkt Potential 1st Shock		-0.02	0.09		0.16	0.16		0.23
		(0.17)	(0.16)		(0.096)*	(0.097)*		(0.099)**
External Mkt Potential 2nd Shock		-0.11	0.13		0.13	0.18		0.19
		(0.16)	(0.16)		(0.09)	(0.090)**		(0.101)*
Constant	0.87	0.82	1.46	0.30	0.87	2.21	16.10	17.70
	(0.100)***	(0.098)***	(0.233)***	(0.39)	(0.274)***	(0.598)***	(14.18)	(14.27)
Observations	4,146	4,146	4,146	4,146	4,146	4,146	4,146	4,146
R-squared	0.34	0.37	0.44	0.16	0.31	0.37	0.21	0.21

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3.III.h: Market Potential as Explanatory Variable Using Productivity and Percentage of Number of Manufacturing Firms – Equation 3.5

Dependent Variable	2SLS			Random Effects			Fixed Effects	
Wi,t / Wsp,t	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)
Internal Mkt Potential	0.05	0.14	0.22	0.16	0.18	0.19	2.27	2.42
	(0.04)	(0.13)	(0.125)*	(0.079)**	(0.12)	(0.12)	(1.62)	(1.203)**
External Mkt Potential	0.12	0.05	0.01	0.05	0.05	0.04	-0.23	-0.35
	(0.028)***	(0.13)	(0.12)	(0.06)	(0.10)	(0.09)	(0.140)*	(0.165)**
Productivity	0.54	0.62	0.59	0.38	0.49	0.41	0.11	0.08
	(0.020)***	(0.026)***	(0.031)***	(0.027)***	(0.041)***	(0.044)***	(0.09)	(0.12)
Subsidies	-0.04	-0.03		-0.03	-0.02		-0.01	-0.01
	(0.005)***	(0.006)***		(0.006)***	(0.006)***		(0.01)	(0.01)
Number of Firms Share	0.01	-0.02	-0.02	0.04	0.00	0.01	0.01	0.00
	(0.01)	(0.007)***	(0.007)***	(0.007)***	(0.01)	(0.01)	(0.02)	(0.01)
Internal Mkt Potential 1st Shock		0.01	0.10		0.00	0.02		-0.19
		(0.16)	(0.15)		(0.10)	(0.10)		(0.103)*
Internal Mkt Potential 2nd Shock		-0.17	-0.06		-0.12	-0.12		-0.18
		(0.15)	(0.14)		(0.10)	(0.09)		(0.104)*
External Mkt Potential 1st Shock		-0.05	-0.14		-0.03	-0.04		0.19
		(0.15)	(0.15)		(0.10)	(0.10)		(0.104)*
External Mkt Potential 2nd Shock		0.13	0.02		0.09	0.09		0.18
		(0.15)	(0.14)		(0.10)	(0.09)		(0.108)*
Constant	0.37	0.27	0.74	0.77	0.57	0.79	12.49	12.63
	(0.085)***	(0.095)***	(0.216)***	(0.182)***	(0.201)***	(0.464)*	(11.15)	(8.02)
Observations	4,145	4,145	4,145	4,145	4,145	4,145	4,145	4,145
R-squared	0.57	0.49	0.55	0.67	0.65	0.69	0.27	0.26

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Annex 3.IV: Further Maps for Chapter 3

Figure 3.IV.a: Map after the Contraction shock Using 1% as Limit

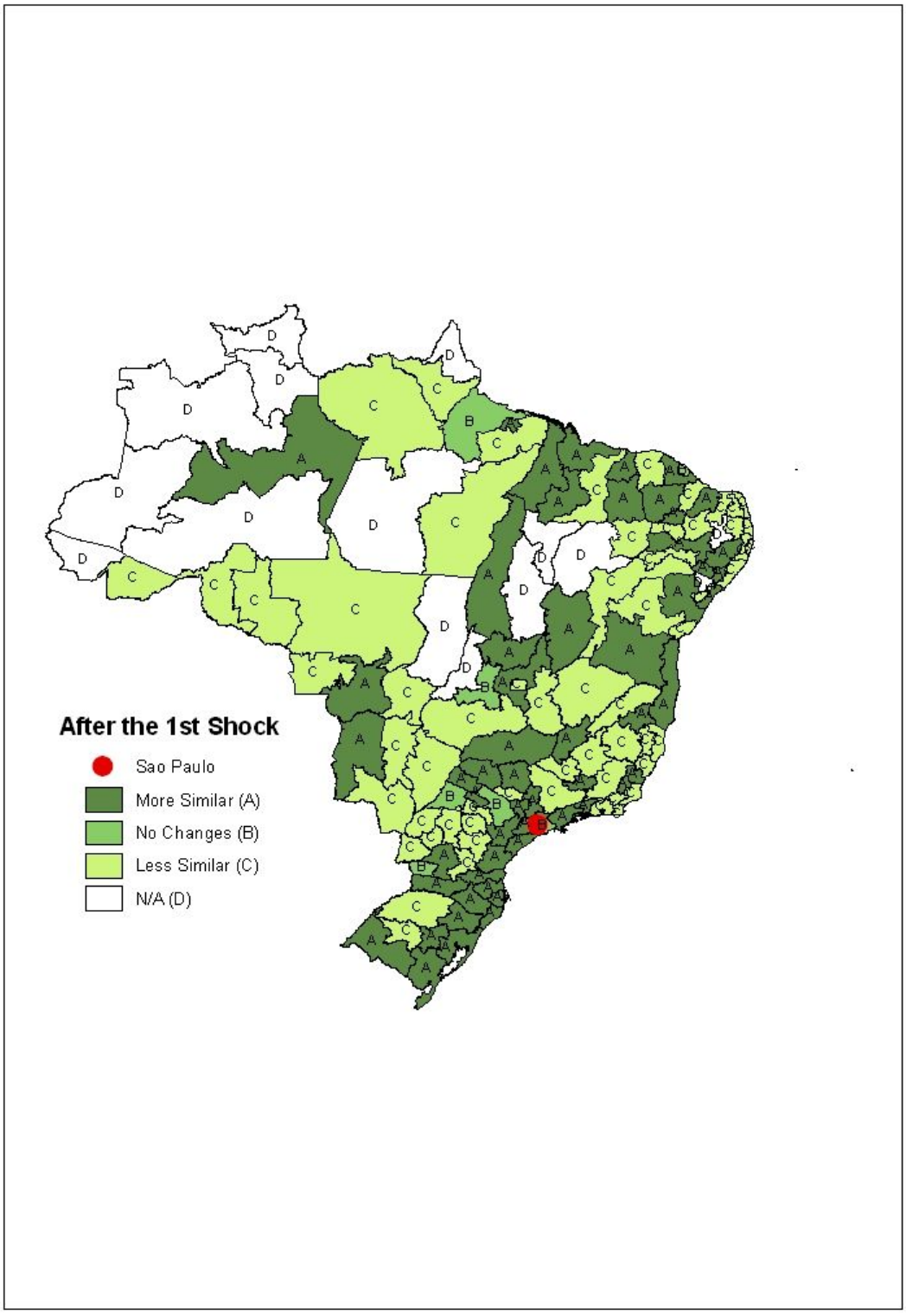


Figure 3.IV.b: Map after the Expansion shock Using 1% as Limit

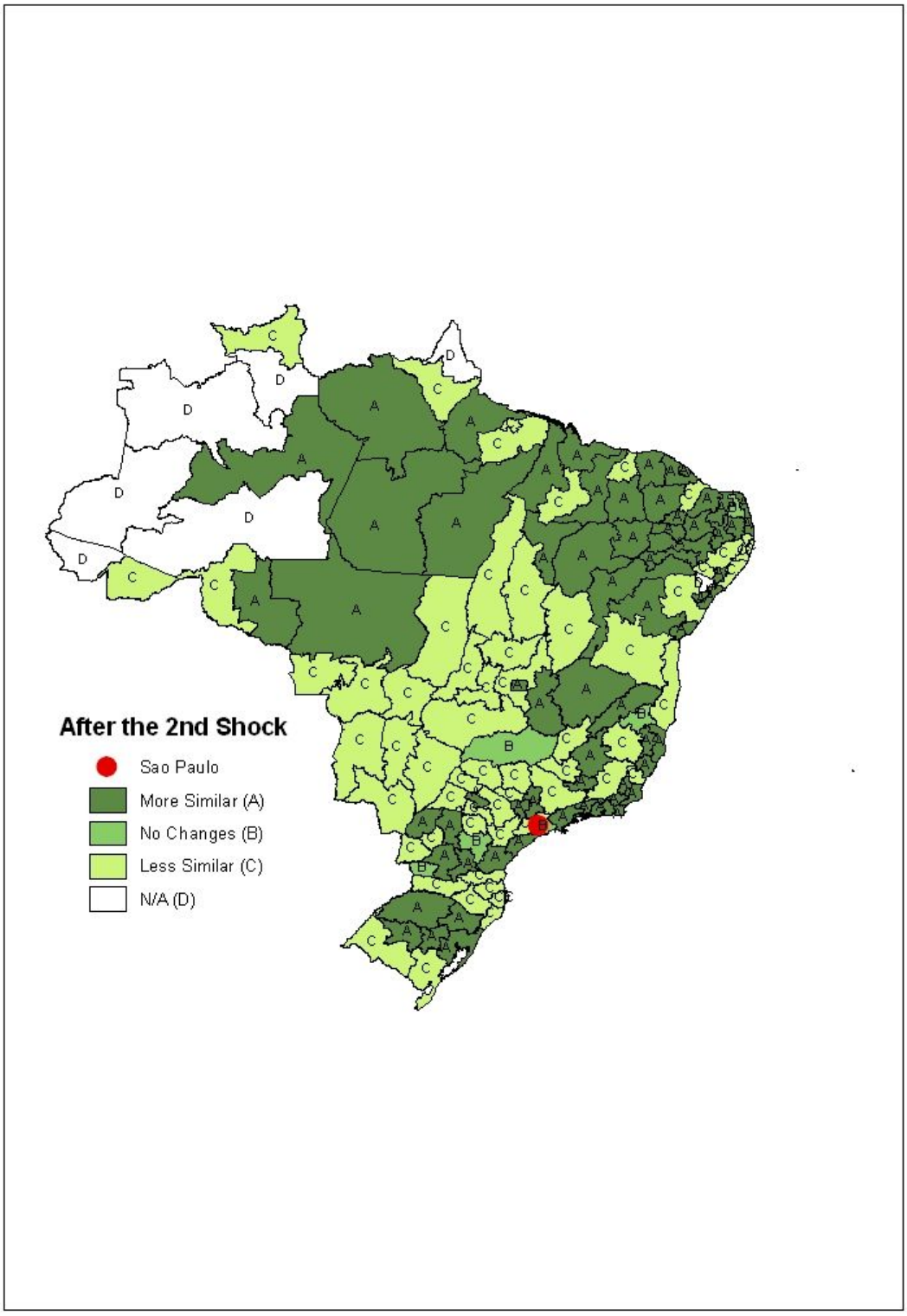


Figure 3.IV.c: Map after the Contraction shock Using 10% as Limit

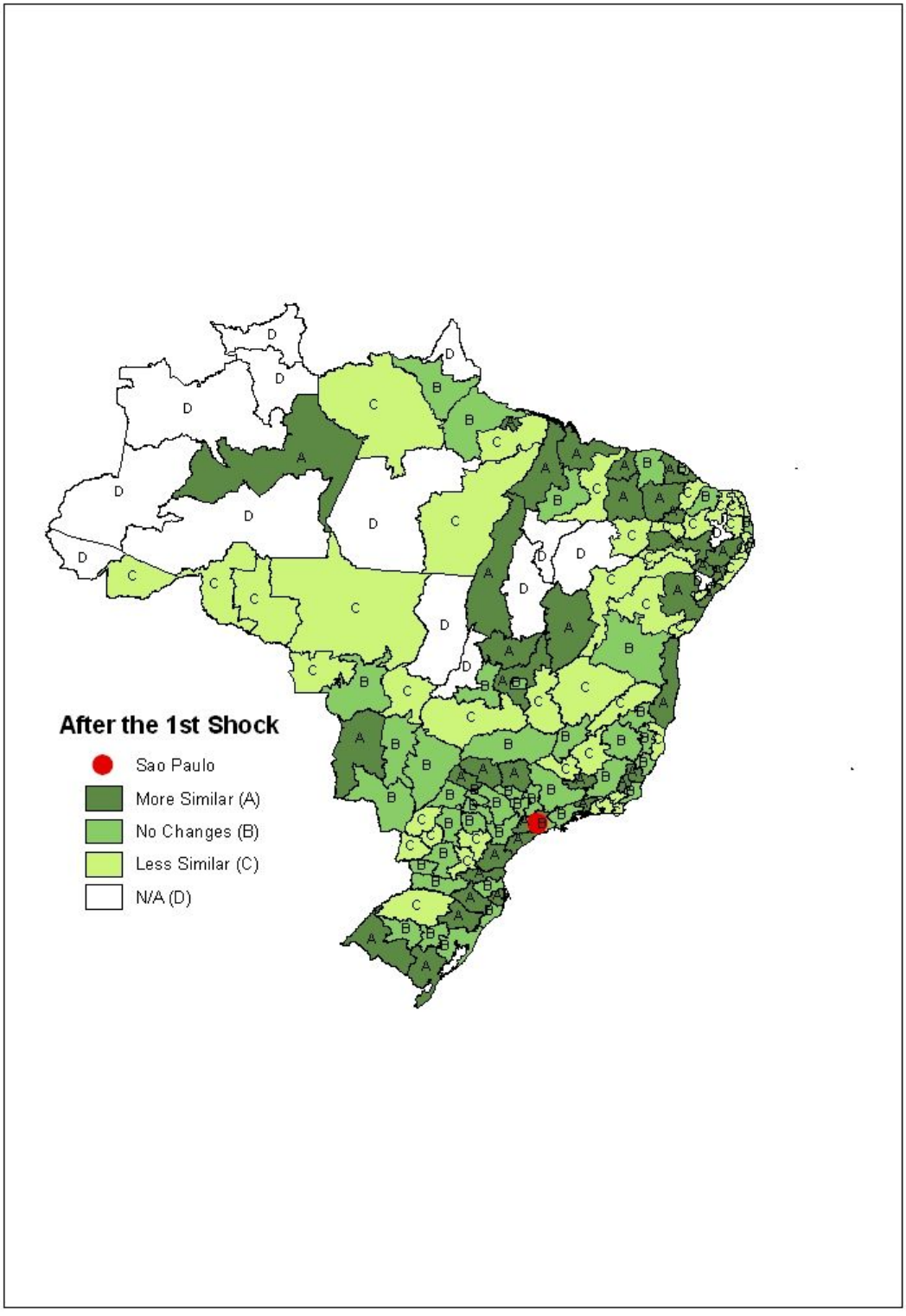
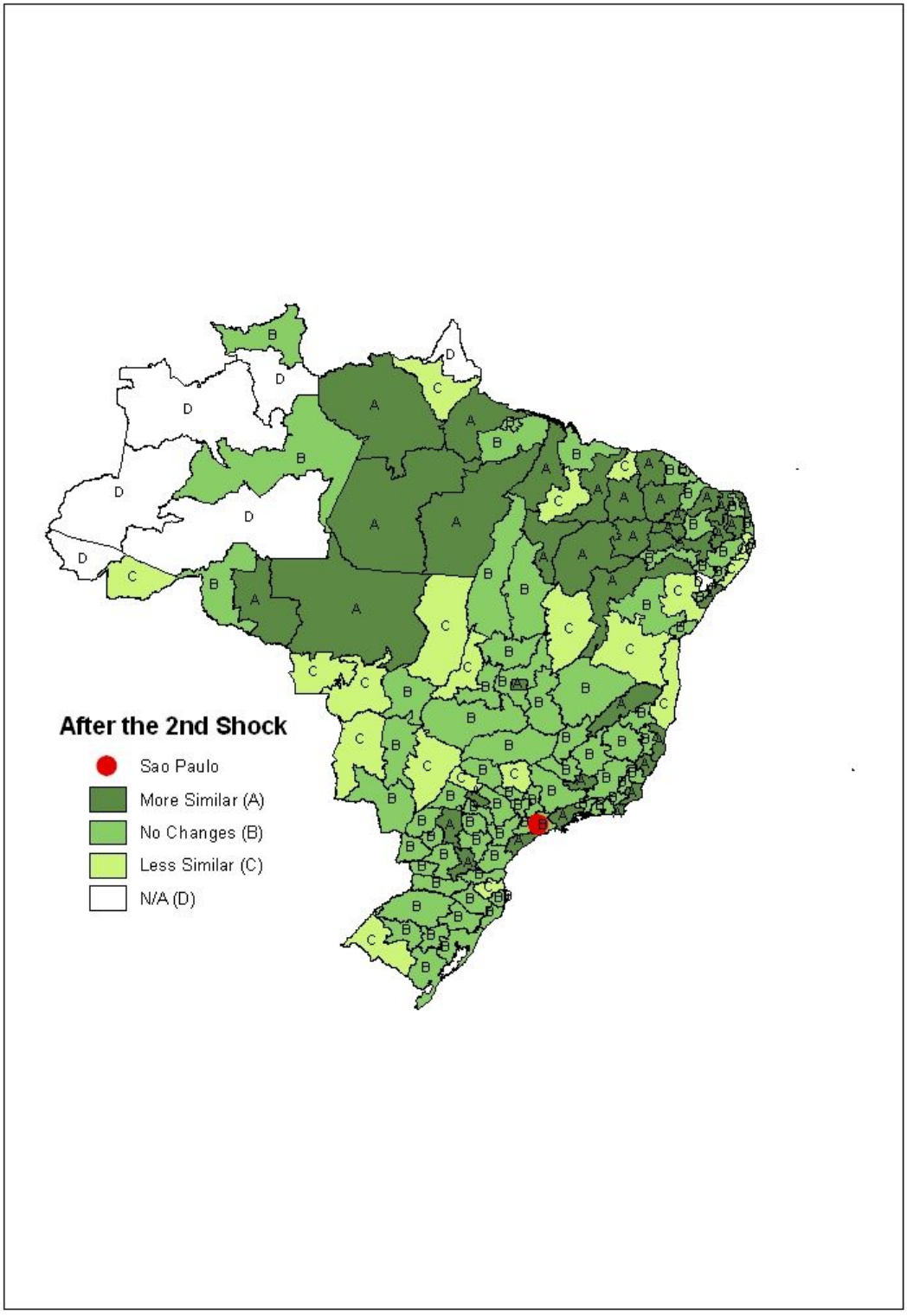


Figure 3.IV.d: Map after the Expansion shock Using 10% as Limit



Appendix 4.I: List of Variables of Chapter 4

Variables which source is PIA from IBGE comprehending information from 1996 to 2003.

1) *Value Added* = total value added;
2) *Number of employees* = total number of employees;
3) *Sector* = sector classification defined by the Brazilian Statistical Institute at the level of CNAE 2, which is similar to SIC 2;
4) *Location* = region where the firm is located (North, Northeast, Middle-West, Southeast and South);
5) *Average salary of employees* = total wages over the number of employees;
6) *Market share* = market share in CNAE 2 sectors in terms of net revenues;
7) *Total revenues* = total value of net revenues, which includes taxes (differently from value added);
8) *Ratio of financial costs to total revenues* = total financial costs over total revenue.

Variables which source is RAIS from Ministry of Labour comprehending information from 1996 to 2003.

1) *Percentage of high skilled workers* = percentage of workers with at least an undergraduate degree;
2) *Educational level of employees* = average number of years spent in schooling by employees;
3) *Age* = number of years in operations until 2003.

Variables which source is SECEX from Ministry of Industrial Development and Foreign Trade jointly with some measures of PIA comprehending information from 1996 to 2003.

1) *Ratio of exports to total revenues* = total exports (SECEX) over total revenues (PIA);
2) *Ratio of imports of capital goods to investment* = total capital goods imports (SECEX) over total investment (PIA);
3) *Ratio of imports of intermediaries to cost of production* = total intermediate imports (SECEX) over total cost of production (PIA).

Variable which source is Brazilian Central Bank from 2000 Foreign Capital Census.

Multinational status = definition of multinational firm by the Central Bank of Brazil for each firm located in Brazil in 2000.

Appendix 4.II: Further Results of Chapter 4

For Tables 4.II.a to 4.II.f presented in this appendix, columns follow this path:
a) Columns (i), (iv) and (vii) show results using only FINAME dummy as control;
b) Columns (ii), (v) and (viii) present outcomes using FINAME, state and manufacturing sector dummy (for simplicity, results of sate and manufacturing dummies are not listed, but they are available upon request);
c) Columns (iii), (vi) and (ix) show results using all controls.

Table 4.II.a: Results for All Projects in Method 1

Dependent Variable	All Regions			Poor Regions			Rich Regions		
Value Added per Worker	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
BNDES	0.47	0.44	0.06	0.63	0.51	0.08	0.44	0.43	0.07
	(0.014)***	(0.014)***	(0.011)***	(0.04)***	(0.04)***	(0.03)**	(0.015)***	(0.014)***	(0.012)***
FINAME	0.56	0.42	0.06	0.60	0.44	0.04	0.54	0.40	0.07
	(0.008)***	(0.008)***	(0.007)***	(0.02)***	(0.02)***	(0.02)**	(0.009)***	(0.009)***	(0.008)***
Firm's age			0.00			-0.02			0.00
			(0.004)			(0.01)			(0.005)
Share of Skilled Worker			0.27			0.51			0.25
			(0.05)***			(0.15)**			(0.06)***
Employee's Wage			0.24			0.17			0.24
			(0.007)***			(0.03)**			(0.007)***
Employee's Education			-0.03			-0.06			-0.02
			(0.012)***			(0.03)**			(0.014)
Number of Employees			-0.64			-0.73			-0.63
			(0.005)***			(0.01)***			(0.005)***
Market Share			-0.16			-1.61			0.09
			(0.37)			(1.69)			(0.39)
Total Revenue			0.74			0.83			0.72
			(0.004)***			(0.009)***			(0.004)***
Financial Status			-0.48			-0.12			-0.58
			(0.034)***			(0.07)*			(0.039)***
Share of Exports			0.07			0.28			0.00
			(0.025)***			(0.05)***			(0.03)
Share of Imported Inputs			-0.18			0.16			-0.22
			(0.055)***			(0.16)			(0.06)***
Share of Imported Capital			0.08			0.13			0.07
			(0.04)**			(0.14)			(0.04)
Multinational Status			0.01			0.00			0.02
			(0.14)			(0.04)			(0.015)
Constant	9.65	9.89	-0.47	9.32	9.60	-0.94	9.72	9.67	-0.55
	(0.004)***	(0.02)***	(0.054)***	(0.012)***	(0.02)***	(0.12)	(0.005)***	(0.011)***	(0.059)***
Observations	189,992	189,992	188,104	30,920	30,920	30,599	159,072	159,072	157,505
R-squared	0.03	0.15	0.42	0.03	0.15	0.45	0.03	0.14	0.41

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4.II.b: Results for Small Projects in Method 1

Dependent Variable	All Regions			Poor Regions			Rich Regions		
Value Added per Worker	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
BNDES	0.41	0.40	0.07	0.57	0.46	0.08	0.38	0.38	0.07
	(0.015)***	(0.015)***	(0.012)***	(0.05)***	(0.05)***	(0.04)**	(0.016)***	(0.015)***	(0.013)***
FINAME	0.57	0.43	0.06	0.62	0.45	0.04	0.55	0.41	0.07
	(0.008)***	(0.008)***	(0.007)***	(0.02)***	(0.02)***	(0.02)**	(0.009)***	(0.009)***	(0.008)***
Firm's age			0.00			-0.02			0.00
			(0.004)			(0.01)			(0.005)
Share of Skilled Worker			0.27			0.51			0.25
			(0.05)***			(0.15)**			(0.06)***
Employee's Wage			0.24			0.17			0.24
			(0.007)***			(0.02)**			(0.007)***
Employee's Education			-0.03			-0.06			-0.02
			(0.012)***			(0.03)**			(0.014)
Number of Employees			-0.64			-0.73			-0.63
			(0.005)***			(0.01)***			(0.005)***
Market Share			-0.07			-1.32			0.18
			(0.38)			(1.69)			(0.39)
Total Revenue			0.74			0.83			0.72
			(0.004)***			(0.009)***			(0.004)***
Financial Status			-0.48			-0.12			-0.58
			(0.034)***			(0.07)*			(0.04)***
Share of Exports			0.07			0.28			0.00
			(0.025)***			(0.05)***			(0.03)
Share of Imported Inputs			-0.18			0.16			-0.22
			(0.055)***			(0.16)			(0.06)***
Share of Imported Capital			0.08			0.14			0.07
			(0.04)**			(0.14)			(0.04)
Multinational Status			0.01			0.00			0.02
			(0.14)			(0.04)			(0.015)
Constant	9.66	9.89	-0.47	9.32	9.60	-0.94	9.73	9.68	-0.55
	(0.004)***	(0.02)***	(0.05)***	(0.012)***	(0.02)***	(0.12)	(0.005)***	(0.012)***	(0.059)***
Observations	189,992	189,992	188,104	30,920	30,920	30,599	159,072	159,072	157,505
R-squared	0.03	0.15	0.42	0.03	0.14	0.45	0.03	0.14	0.41

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4.II.c: Results for Large Projects in Method 1

Dependent Variable	All Regions			Poor Regions			Rich Regions		
Value Added per Worker	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
BNDES	1.08	0.97	0.06	1.10	0.94	0.09	1.09	0.98	0.06
	(0.04)***	(0.04)***	(0.03)*	(0.10)***	(0.10)***	-0.08	(0.04)***	(0.04)***	(-0.04)
FINAME	0.59	0.45	0.07	0.64	0.47	0.04	0.57	0.44	0.08
	(0.008)***	(0.008)***	(0.007)***	(0.02)***	(0.02)***	(0.02)**	(0.009)***	(0.009)***	(0.007)***
Firm's age			0.00			-0.01			0.00
			(0.004)			(0.01)			(0.005)
Share of Skilled Worker			0.27			0.52			0.26
			(0.05)***			(0.15)**			(0.06)***
Employee's Wage			0.23			0.17			0.24
			(0.007)***			(0.02)**			(0.007)***
Employee's Education			-0.03			-0.06			-0.01
			(0.012)**			(0.03)**			(0.014)
Number of Employees			-0.64			-0.73			-0.62
			(0.005)***			(0.01)***			(0.005)***
Market Share			-0.17			-1.64			0.09
			(0.38)			(1.72)			(0.39)
Total Revenue			0.74			0.83			0.72
			(0.004)***			(0.009)***			(0.004)***
Financial Status			-0.48			-0.12			-0.58
			(0.03)***			(0.07)*			(0.04)***
Share of Exports			0.07			0.28			0.00
			(0.02)***			(0.05)***			(0.03)
Share of Imported Inputs			-0.18			0.16			-0.22
			(0.06)***			(0.16)			(0.06)***
Share of Imported Capital			0.08			0.14			0.07
			(0.04)*			(0.14)			(0.04)
Multinational Status			0.01			0.00			0.02
			(0.014)			(0.04)			(0.015)
Constant	9.66	9.89	-0.48	9.33	9.61	-0.95	9.73	9.68	-0.56
	(0.004)***	(0.02)***	(0.05)***	(0.012)***	(0.02)***	(0.12)	(0.005)***	(0.012)***	(0.059)***
Observations	189,992	189,992	188,104	30,920	30,920	30,599	159,072	159,072	157,505
R-squared	0.03	0.15	0.42	0.03	0.14	0.45	0.03	0.14	0.41

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4.II.d: Results for All Projects in Method 2

Dependent Variable	All Regions			Poor Regions			Rich Regions		
Value Added per Worker	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
BNDES	0.13	0.12	0.05	0.22	0.20	0.05	0.11	0.12	0.06
	(0.017)***	(0.016)***	(0.013)***	(0.05)***	(0.05)***	(0.04)	(0.02)***	(0.02)***	(0.014)***
Ever BNDES Automatic	0.49	0.35	0.02	0.33	0.26	0.03	0.33	0.34	0.03
	(0.008)***	(0.013)***	(0.011)	(0.04)***	(0.04)***	(0.03)	(0.015)***	(0.014)***	(0.012)**
Ever FINEM	1.36	1.18	0.02	1.53	1.23	0.11	1.35	1.15	-0.14
	(0.04)***	(0.03)***	(0.03)	(0.08)***	(0.08)***	(0.06)*	(0.04)***	(0.04)***	(0.03)
FINAME	0.34	0.35	0.06	0.52	0.37	0.03	0.48	0.34	0.07
	(0.014)***	(0.008)***	(0.007)***	(0.02)***	(0.02)***	(0.07)*	(0.009)***	(0.009)***	(0.008)***
Firm's age			0.00			-0.01			0.00
			(0.004)			(0.01)			(0.005)
Share of Skilled Worker			0.27			0.49			0.25
			(0.05)***			(0.15)***			(0.06)***
Employee's Wage			0.24			0.17			0.24
			(0.007)***			(0.02)***			(0.007)***
Employee's Education			-0.03			-0.06			-0.02
			(0.012)***			(0.03)**			(0.014)
Number of Employees			-0.64			-0.73			-0.63
			(0.005)***			(0.01)***			(0.005)***
Market Share			-0.17			-1.93			0.15
			(0.38)			(1.71)			(0.39)
Total Revenue			0.74			0.83			0.72
			(0.004)***			(0.009)***			(0.004)***
Financial Status			-0.48			-0.12			-0.58
			(0.034)***			(0.07)*			(0.04)***
Share of Exports			0.07			0.28			0.00
			(0.025)***			(0.05)***			(0.03)
Share of Imported Inputs			-0.18			0.16			-0.22
			(0.055)***			(0.16)			(0.06)***
Share of Imported Capital			0.05			0.13			0.07
			(0.03)**			(0.14)			(0.04)
Multinational Status			0.01			0.00			0.02
			(0.014)			(0.04)			(0.015)
Constant	9.64	9.88	-0.46	9.31	9.59	-0.91	9.71	9.65	-0.54
	(0.005)***	(0.02)***	(0.06)***	(0.012)***	(0.02)***	(0.12)	(0.005)***	(0.011)***	(0.059)***
Observations	189,992	189,992	188,104	30,920	30,920	30,599	159,072	159,072	157,505
R-squared	0.04	0.16	0.42	0.05	0.15	0.45	0.04	0.15	0.41

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4.II.e: Results for Small Projects in Method 2

Dependent Variable	All Regions			Poor Regions			Rich Regions		
Value Added per Worker	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
BNDES	0.13	0.13	0.06	0.26	0.22	0.06	0.11	0.12	0.06
	(0.02)***	(0.016)***	(0.014)***	(0.05)***	(0.05)***	(0.04)	(0.02)***	(0.02)***	(0.014)***
Ever BNDES Automatic	0.34	0.35	0.02	0.32	0.25	0.03	0.33	0.34	0.03
	(0.008)***	(0.013)***	(0.011)	(0.04)***	(0.04)***	(0.03)	(0.015)***	(0.015)***	(0.012)**
Ever FINEM	1.42	1.24	0.05	1.60	1.30	0.12	1.40	1.20	0.01
	(0.04)***	(0.03)***	(0.03)	(0.08)***	(0.07)***	(0.06)**	(0.04)***	(0.04)***	(0.03)
FINAME	0.49	0.35	0.06	0.52	0.37	0.03	0.48	0.34	0.07
	(0.009)***	(0.008)***	(0.007)***	(0.02)***	(0.02)***	(0.02)*	(0.009)***	(0.009)***	(0.008)***
Firm's age			0.00			-0.01			0.00
			(0.005)			(0.01)			(0.005)
Share of Skilled Worker			0.27			0.49			0.25
			(0.05)***			(0.16)***			(0.06)***
Employee's Wage			0.24			0.17			0.24
			(0.007)***			(0.02)***			(0.007)***
Employee's Education			-0.03			-0.06			-0.02
			(0.012)***			(0.03)**			(0.015)
Number of Employees			-0.64			-0.73			-0.63
			(0.005)***			(0.01)***			(0.005)***
Market Share			-0.14			-1.81			0.18
			(0.38)			(1.71)			(0.39)
Total Revenue			0.74			0.83			0.72
			(0.004)***			(0.009)***			(0.004)***
Financial Status			-0.48			-0.12			-0.58
			(0.034)***			(0.07)*			(0.04)***
Share of Exports			0.07			0.28			0.00
			(0.025)***			(0.05)***			(0.03)
Share of Imported Inputs			-0.18			0.16			-0.22
			(0.055)***			(0.16)			(0.06)***
Share of Imported Capital			0.05			0.13			0.07
			(0.03)**			(0.14)			(0.04)
Multinational Status			0.01			0.00			0.02
			(0.014)			(0.04)			(0.015)
Constant	9.64	9.88	-0.46	9.31	9.59	-0.91	9.71	9.65	-0.54
	(0.005)***	(0.02)***	(0.06)***	(0.012)***	(0.02)***	(0.12)	(0.005)***	(0.011)***	(0.059)***
Observations	189,992	189,992	188,104	30,920	30,920	30,599	159,072	159,072	157,505
R-squared	0.04	0.16	0.42	0.05	0.15	0.45	0.04	0.15	0.41

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4.II.f: Results for Large Projects in Method 2

Dependent Variable	All Regions			Poor Regions			Rich Regions		
Value Added per Worker	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
BNDES	0.09	0.11	0.04	0.03	0.12	0.01	0.09	0.13	0.06
	(0.05)***	(0.05)**	(0.04)	(0.11)	(0.11)	(0.09)	(0.06)*	(0.05)**	(0.04)
Ever BNDES Automatic	0.39	0.40	0.04	0.41	0.33	0.05	0.38	0.39	0.05
	(0.013)***	(0.012)***	(0.010)***	(0.04)***	(0.04)***	(0.03)	(0.013)***	(0.013)***	(0.011)**
Ever FINEM	1.39	1.20	0.03	1.62	1.28	0.13	1.37	1.15	-0.01
	(0.04)***	(0.04)***	(0.03)	(0.09)***	(0.08)***	(0.07)*	(0.05)***	(0.05)***	(0.04)
FINAME	0.49	0.36	0.06	0.52	0.37	0.03	0.48	0.34	0.07
	(0.009)***	(0.008)***	(0.007)***	(0.02)***	(0.02)***	(0.02)*	(0.009)***	(0.009)***	(0.008)***
Firm's age			0.00			-0.01			0.00
			(0.005)			(0.01)			(0.005)
Share of Skilled Worker			0.27			0.50			0.25
			(0.05)***			(0.16)***			(0.06)***
Employee's Wage			0.23			0.17			0.24
			(0.007)***			(0.02)***			(0.007)***
Employee's Education			-0.03			-0.06			-0.01
			(0.012)***			(0.03)**			(0.015)
Number of Employees			-0.64			-0.73			-0.63
			(0.005)***			(0.01)***			(0.005)***
Market Share			-0.17			-1.85			0.14
			(0.38)			(1.72)			(0.39)
Total Revenue			0.74			0.83			0.72
			(0.004)***			(0.009)***			(0.004)***
Financial Status			-0.48			-0.13			-0.58
			(0.034)***			(0.07)*			(0.04)***
Share of Exports			0.07			0.28			0.00
			(0.025)***			(0.05)***			(0.03)
Share of Imported Inputs			-0.18			0.16			-0.22
			(0.06)***			(0.16)			(0.06)***
Share of Imported Capital			0.05			0.13			0.07
			(0.03)**			(0.14)			(0.04)
Multinational Status			0.01			0.00			0.02
			(0.014)			(0.04)			(0.015)
Constant	9.64	9.88	-0.46	9.31	9.59	-0.91	9.71	9.65	-0.54
	(0.005)***	(0.02)***	(0.06)***	(0.012)***	(0.02)***	(0.12)	(0.005)***	(0.011)***	(0.059)***
Observations	189,992	189,992	188,104	30,920	30,920	30,599	159,072	159,072	157,505
R-squared	0.04	0.16	0.42	0.05	0.15	0.45	0.04	0.15	0.41

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

For Tables 4.II.g to 4.II.l presented in this appendix, columns follow this path:

a) Columns (i), (iii) and (v) show results without any control;

b) Columns (ii), (iv) and (vi) show results using all controls.

Table 4.II.g: Results for All Projects in Method 3

Dependent Variable	All Regions		Poor Regions		Rich Regions	
Value Added per Worker	(i)	(ii)	(iii)	(iv)	(v)	(vi)
BNDES	-0.01	-0.01	-0.04	-0.06	-0.01	0.00
	(0.013)***	(0.012)***	(0.03)	(0.03)*	(0.014)	(0.012)
Firm's age		-0.18		-0.05		-0.21
		(0.02)***		(0.05)		(0.03)***
Share of Skilled Worker		-0.16		-0.11		-0.15
		(0.09)*		(0.22)		(0.10)***
Employee's Wage		0.05		0.01		0.06
		(0.013)***		(0.03)		(0.015)***
Employee's Education		-0.05		-0.03		-0.05
		(0.09)*		(0.04)		(0.03)
Number of Employees		-0.85		-0.90		-0.83
		(0.012)***		(0.03)***		(0.01)***
Market Share		-0.95		-4.91		-0.89
		(0.89)		(2.42)**		(0.71)
Total Revenue		1.03		1.07		1.02
		(0.011)***		(0.03)***		(0.012)***
Financial Status		-0.05		0.25		-0.15
		(0.11)		(0.13)*		(0.14)
Share of Exports		-0.56		-0.29		-0.63
		(0.09)		(0.12)**		(0.10)***
Share of Imported Inputs		0.01		0.61		-0.02
		(0.06)		(0.11)***		(0.08)
Share of Imported Capital		0.04		-0.03		0.05
		(0.04)		(0.08)		(0.04)
Constant	9.85	-2.11	9.51	9.31	9.92	-1.88
	(0.003)***	(0.33)***	(0.007)***	(0.012)***	(0.03)***	(0.22)
Observations	189,992	188,104	30,920	30,599	159,072	157,505
R-squared	0.01	0.39	0.01	0.41	0.01	0.38

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4.II.h: Results for Small Projects in Method 3

Dependent Variable	All Regions		Poor Regions		Rich Regions	
Value Added per Worker	(i)	(ii)	(iii)	(iv)	(v)	(vi)
BNDES	-0.03	-0.01	-0.07	-0.06	-0.02	0.00
	(0.013)*	(0.014)***	(0.04)*	(0.03)	(0.016)	(0.015)
Firm's age		-0.18		-0.05		-0.22
		(0.02)***		(0.05)		(0.03)***
Share of Skilled Worker		-0.16		-0.11		-0.15
		(0.09)*		(0.22)		(0.10)***
Employee's Wage		0.05		0.01		0.06
		(0.013)***		(0.03)		(0.015)***
Employee's Education		-0.05		-0.03		-0.05
		(0.03)*		(0.04)		(0.03)
Number of Employees		-0.85		-0.90		-0.83
		(0.012)***		(0.03)***		(0.013)***
Market Share		-0.97		-5.17		-0.89
		(0.89)		(2.40)**		(0.71)
Total Revenue		1.03		1.07		1.02
		(0.011)***		(0.03)***		(0.012)***
Financial Status		-0.05		0.25		-0.15
		(0.11)		(0.13)*		(0.14)
Share of Exports		-0.56		-0.29		-0.63
		(0.09)***		(0.12)**		(0.10)***
Share of Imported Inputs		0.01		0.61		-0.02
		(0.06)		(0.11)***		(0.08)
Share of Imported Capital		0.04		-0.03		0.05
		(0.04)		(0.08)		(0.04)
Constant	9.85	-2.11	9.51	-2.79	9.92	-1.88
	(0.003)***	(0.33)***	(0.007)***	(0.87)***	(0.03)***	(0.22)
Observations	189,992	188,104	30,920	30,599	159,072	157,505
R-squared	0.01	0.39	0.01	0.41	0.01	0.38

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4.II.i: Results for Large Projects in Method 3

Dependent Variable	All Regions		Poor Regions		Rich Regions	
Value Added per Worker	(i)	(ii)	(iii)	(iv)	(v)	(vi)
BNDES	0.09	0.00	0.06	-0.06	0.09	0.01
	(0.02)***	(0.016)	(0.06)	(0.04)	(0.03)***	(0.02)
Firm's age		-0.18		-0.05		-0.22
		(0.02)***		(0.05)		(0.03)***
Share of Skilled Worker		-0.16		-0.12		-0.15
		(0.09)*		(0.22)		(0.10)
Employee's Wage		0.05		0.02		0.06
		(0.013)***		(0.03)		(0.015)***
Employee's Education		-0.05		-0.04		-0.05
		(0.03)*		(0.05)		(0.03)
Number of Employees		-0.85		-0.90		-0.84
		(0.012)***		(0.03)***		(0.013)***
Market Share		-0.97		-4.94		-0.92
		(0.90)		(2.45)**		(0.71)
Total Revenue		1.03		1.07		1.02
		(0.011)***		(0.03)***		(0.012)***
Financial Status		-0.05		0.25		-0.15
		(0.11)		(0.13)*		(0.14)
Share of Exports		-0.56		-0.29		-0.63
		(0.09)***		(0.12)**		(0.11)***
Share of Imported Inputs		0.01		0.61		-0.02
		(0.06)		(0.11)***		(0.08)
Share of Imported Capital		0.04		-0.03		0.05
		(0.04)		(0.08)		(0.04)
Constant	9.85	-2.11	9.51	-2.79	9.92	-1.88
	(0.003)***	(0.34)***	(0.007)***	(0.87)***	(0.003)***	(0.22)
Observations	189,992	188,104	30,920	30,599	159,072	157,505
R-squared	0.01	0.39	0.01	0.41	0.01	0.38

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4.II.j: Results for All Projects in Method 4

Dependent Variable	All Regions		Poor Regions		Rich Regions	
Value Added per Worker	(i)	(ii)	(iii)	(iv)	(v)	(vi)
BNDES	0.00	0.01	-0.01	0.00	0.00	0.01
	(0.014)	(0.013)	(0.04)	(0.04)	(0.014)	(0.014)
Firm's age		-0.03		-0.05		-0.05
		(0.06)		(0.05)		(0.07)
Share of Skilled Worker		0.10		0.36		0.06
		(0.13)		(0.31)		(0.13)***
Employee's Wage		-0.06		-0.01		-0.07
		(0.02)***		(0.05)		(0.02)***
Employee's Education		-0.01		-0.02		-0.01
		(0.03)		(0.06)		(0.03)
Number of Employees		-0.91		-0.91		-0.91
		(0.01)***		(0.02)***		(0.01)***
Market Share		-0.97		-2.58		-0.61
		(2.36)		(7.25)		(2.49)
Total Revenue		1.08		1.08		1.08
		(0.01)***		(0.02)***		(0.011)***
Financial Status		0.20		0.35		0.16
		(0.05)***		(0.11)***		(0.06)
Share of Exports		-0.75		-0.45		-0.83
		(0.06)***		(0.14)***		(0.07)***
Share of Imported Inputs		0.06		0.73		0.03
		(0.06)		(0.29)***		(0.06)
Share of Imported Capital		-0.02		-0.09		-0.02
		(0.06)		(0.17)		(0.07)
Multinational Status		-0.01		-0.01		-0.01
		(0.16)		(0.05)		(0.02)
Constant	-0.02	-0.03	-0.02	-0.03	-0.02	-0.05
	(0.004)***	(0.03)***	(0.011)**	(0.02)	(0.004)***	(0.014)
Observations	138,434	136,457	21,636	21,359	116,798	115,098
R-squared	0.00	0.10	0.00	0.11	0.00	0.09

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4.II.k: Results for Small Projects in Method 4

Dependent Variable	All Regions		Poor Regions		Rich Regions	
Value Added per Worker	(i)	(ii)	(iii)	(iv)	(v)	(vi)
BNDES	0.00	0.01	-0.01	0.01	0.00	0.01
	(0.015)	(0.014)	(0.05)	(0.04)	(0.015)	(0.015)
Firm's age		-0.03		0.04		-0.05
		(0.06)		(0.14)		(0.07)
Share of Skilled Worker		0.10		0.36		0.06
		(0.13)		(0.31)		(0.13)***
Employee's Wage		-0.06		-0.01		-0.07
		(0.02)***		(0.05)		(0.02)***
Employee's Education		-0.01		-0.02		-0.01
		(0.03)		(0.06)		(0.03)
Number of Employees		-0.91		-0.91		-0.91
		(0.01)***		(0.02)***		(0.01)***
Market Share		-0.95		-2.58		-0.59
		(2.36)		(7.25)		(2.49)
Total Revenue		1.08		1.09		1.08
		(0.01)***		(0.02)***		(0.011)***
Financial Status		0.20		0.35		0.16
		(0.05)***		(0.11)***		(0.06)
Share of Exports		-0.75		-0.45		-0.84
		(0.06)***		(0.14)***		(0.07)***
Share of Imported Inputs		0.06		0.74		0.03
		(0.06)		(0.29)***		(0.06)
Share of Imported Capital		-0.02		-0.09		-0.02
		(0.06)		(0.17)		(0.07)
Multinational Status		-0.01		-0.01		-0.01
		(0.16)		(0.05)		(0.02)
Constant	-0.02	-0.03	-0.02	-0.04	-0.02	-0.05
	(0.004)***	(0.03)***	(0.011)**	(0.03)	(0.005)***	(0.014)
Observations	138,434	136,457	21,636	21,359	116,798	115,098
R-squared	0.00	0.10	0.00	0.11	0.00	0.09

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4.II.l: Results for Large Projects in Method 4

Dependent Variable	All Regions		Poor Regions		Rich Regions	
Value Added per Worker	(i)	(ii)	(iii)	(iv)	(v)	(vi)
BNDES	0.01	0.00	-0.01	-0.02	0.02	0.01
	(0.04)	(0.04)	(0.10)	(0.09)	(0.04)	(0.04)
Firm's age		-0.03		0.04		-0.05
		(0.06)		(0.15)		(0.07)
Share of Skilled Worker		0.10		0.36		0.06
		(0.13)		(0.31)		(0.13)***
Employee's Wage		-0.06		-0.01		-0.07
		(0.02)***		(0.05)		(0.02)
Employee's Education		-0.01		-0.02		-0.01
		(0.03)		(0.06)		(0.03)
Number of Employees		-0.91		-0.91		-0.91
		(0.01)***		(0.02)***		(0.01)***
Market Share		-0.95		-2.49		-0.61
		(2.36)		(7.27)		(2.49)
Total Revenue		1.08		1.09		1.08
		(0.01)***		(0.02)***		(0.011)***
Financial Status		0.20		0.35		0.16
		(0.05)***		(0.11)***		(0.06)***
Share of Exports		-0.75		-0.45		-0.84
		(0.06)***		(0.14)***		(0.07)***
Share of Imported Inputs		0.06		0.74		0.03
		(0.06)		(0.29)***		(0.06)
Share of Imported Capital		-0.02		-0.09		-0.02
		(0.06)		(0.17)		(0.07)
Multinational Status		-0.01		-0.01		-0.01
		(0.02)		(0.05)		(0.02)
Constant	-0.02	-0.03	-0.02	-0.03	-0.02	-0.05
	(0.004)***	(0.03)	(0.011)**	(0.03)	(0.005)***	(0.014)
Observations	138,434	136,457	21,636	21,359	116,798	115,098
R-squared	0.00	0.10	0.00	0.11	0.00	0.09

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

In Tables 4.II.m, columns follow this path: a) Columns (i), (iii) and (v) show results using only FINAME and FINEM dummy as controls; b) Columns (ii), (iv) and (vi) present outcomes using all controls.

Table 4.II.m: Results for Small Projects in Method 5

Dependent Variable	All Regions		Poor Regions		Rich Regions	
Value Added per Worker	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Policy effect in 97	-0.07	-0.07	0.04	0.10	-0.08	-0.09
	(0.10)	(0.07)	(0.28)	(0.13)	(0.11)	(0.11)
Policy effect in 98	0.09	0.07	-0.09	-0.07	0.10	0.09
	(0.10)	(0.08)	(0.29)	(0.15)	(0.11)	(0.11)
Policy effect in 99	0.06	0.06	0.06	-0.01	0.06	0.06
	(0.09)	(0.06)	(0.32)	(0.19)	(0.10)	(0.12)
Policy effect in 00	-0.11	-0.14	0.13	0.03	-0.13	-0.14
	(0.11)	(0.08)*	(0.33)	(0.22)	(0.11)	(0.12)
Policy effect in 01	-0.03	0.02	-0.02	0.04	-0.03	0.00
	(0.12)	(0.10)	(0.32)	(0.17)	(0.13)	(0.12)
Policy effect in 02	-0.07	-0.07	-0.03	-0.10	-0.07	-0.06
	(0.11)	(0.09)	(0.33)	(0.17)	(0.12)	(0.12)
Policy effect in 03	0.08	0.08	0.08	-0.13	0.08	0.07
	(0.10)	(0.07)	(0.35)	(0.21)	(0.11)	(0.13)
Finame	0.46	0.05	0.22	-0.15	0.49	0.09
	(0.04)***	(0.03)*	(0.09)***	(0.06)**	(0.04)***	(0.02)***
Age		0.03		0.00		0.04
		(0.02)		(0.07)		(0.02)*
Skill		0.22		0.26		0.37
		(0.17)		(0.50)		(0.17)**
Wage		0.38		0.23		0.30
		(0.03)***		(0.05)***		(0.03)
Schooling		0.01		0.20		-0.06
		(0.05)		(0.14)		(0.06)
Number of employees		-0.63		-0.63		-0.63
		(0.02)***		(0.03)***		(0.02)***
Market Share		1.29		-11.22		-1.82
		(1.02)		(5.34)**		(1.72)
Revenues		0.70		0.78		0.73
		(0.02)***		(0.03)***		(0.02)***
Financial Status		-0.25		-1.44		-0.18
		(0.25)		(1.24)		(0.15)
Export		0.35		0.52		0.13
		(0.05)***		(0.12)***		(0.07)***
Intermediaries import		-0.76		0.79		-1.59
		(0.60)***		(0.28)***		(0.20)***
Capital import		0.04		0.48		0.00
		(0.07)		(0.14)***		(0.11)
FINEM	1.02	-0.01	1.16	0.48	1.01	0.21
	(0.04)***	(0.02)	(0.17)***	(0.14)***	(0.04)***	(0.05)
Number of Obs.	12,172	12,132	1,355	1,352	10,817	10,780
R-squared	0.06	0.50	0.07	0.69	0.06	0.50

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

In Tables 4.II.n, columns follow this path: a) Columns (i), (iii) and (v) show results using only FINAME and Automatic BNDES dummy as controls; b) Columns (ii), (iv) and (vi) present outcomes using all controls.

Table 4.II.n: Results for Large Projects in Method 5

Dependent Variable	All Regions		Poor Regions		Rich Regions	
Value Added per Worker	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Policy effect in 97	0.07	0.00	-0.05	-0.49	0.08	0.08
	(0.30)	(0.09)	(1.15)	(0.11)*	(0.30)	(0.38)
Policy effect in 98	-0.03	-0.06	0.30	0.16	-0.07	-0.04
	(0.30)	(0.11)	(0.86)	(0.13)	(0.32)	(0.38)
Policy effect in 99	0.08	0.03	-0.47	-0.08	0.14	0.14
	(0.35)	(0.21)	(0.82)	(0.19)	(0.37)	(0.42)
Policy effect in 00	-0.11	-0.06	-2.14	-1.74	0.20	0.17
	(0.41)	(0.36)	(1.17)*	(1.32)	(0.38)	(0.44)
Policy effect in 01	0.27	0.19	1.82	1.69	0.02	0.00
	(0.38)	(0.31)	(1.13)	(1.31)	(0.35)	(0.45)
Policy effect in 02	-0.31	-0.16	0.29	0.08	-0.40	-0.28
	(0.37)	(0.14)	(1.05)	(0.32)	(0.39)	(0.45)
Policy effect in 03	0.12	0.00	-0.11	-0.13	0.15	-0.04
	(0.37)	(0.14)	(1.17)	(0.32)	(0.38)	(0.44)
Finame	0.01	0.02	0.42	-0.01	-0.02	0.00
	(0.07)	(0.04)	(0.16)***	(0.06)	(0.06)	(0.02)
Age		0.07		-0.01		0.07
		(0.05)**		(0.08)		(0.03)**
Skill		0.19		1.04		-0.08
		(0.23)		(0.37)***		(0.18)
Wage		0.11		-0.23		0.33
		(0.07)***		(0.17)		(0.03)
Schooling		0.11		-0.12		0.18
		(0.07)		(0.11)		(0.07)**
Number of employees		-0.73		-0.84		-0.61
		(0.05)***		(0.09)***		(0.02)***
Market Share		-0.99		-16.05		1.14
		(1.10)		(10.07)		(1.66)
Revenues		0.81		1.14		0.70
		(0.04)***		(0.13)***		(0.02)***
Financial Status		-0.03		-1.17		-0.80
		(0.14)		(0.55)**		(0.11)***
Export		-0.44		-0.24		-0.55
		(0.14)***		(0.23)		(0.08)***
Intermediaries import		0.04		1.05		0.05
		(0.06)		(0.28)***		(0.07)
Capital import		0.29		0.32		0.21
		(0.17)*		(0.25)		(0.11)**
Automatic BNDES	0.07	0.01	-0.24	-0.18	0.10	0.07
	(0.08)	(0.04)	(0.18)	(0.06)***	(0.08)	(0.03)**
Number of Obs.	11,537	11,435	1,269	1,265	10,268	10,170
R-squared	0.01	0.64	0.14	0.74	0.01	0.39

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix III: Kernel Matching

Kernel matching is a method used to compare non-treated observations to treated ones using Kernel density. The first step is estimating a probit model with the variables described by the researcher, which calculates how important each characteristic is to explain being granted a loan in BNDES or not. For this probit model, all controls used for estimation plus productivity are used as explanatory variables in this probit. Therefore, the equation used for the probit is the following:

$$\Pr(BNDES_{i,t}) = \alpha_0 + \alpha_1 produtivity_{i,t-1} + \alpha_2 Empl_{i,t-1} + ... + \varepsilon_{it}$$

where $\Pr(BNDES_{i,t})$ is the probability of getting BNDES support of firm *i* at time *t*, which takes value one if a firm got it and zero otherwise; $produtivity_{i,t-1}$ is the productivity of firm *i* at time *t-1*; $Empl_{i,t-1}$ is the employment size of firm *i* at time *t-1*; and other controls.

After having the coefficients from the probit model, a weight is given to each firm through a kernel density. Kernel [*K(u)*] is a density function which satisfies two assumptions:

1) $\int_{-\infty}^{+\infty} K(u)du = 1$ as for any density function;

2) *K(u)=K(-u)* for all values of *u*

Therefore, any non-treated firm receives a weight which represents how similar they are from the treated one, according to the characteristics described by the researcher.

Bibliography

Amadeo, E. J. and G. Gonzaga (1996). “Salario, Produtividade e Cambio: Uma Analise do Custo Unitario na Indústria Brasileira (1985-1995).” Departamento de Economia, PUC-RJ. Rio de Janeiro. Brazil.

Andrade, T. A., A. M. S. P. Santos and R. V. Serra (2000). “Fluxos migratórios nas cidades médias e regiões metropolitanas brasileiras: a experiência do período 1980/96.” XII Encontro Nacional de Estudos Populacionais, Caxambu (MG), Associação Brasileira de Estudos Populacionais. Brazil.

Azzoni, C. R. and L. M. S. Servo (2002). “Education, cost of living and regional wage inequality in Brazil.” Papers in Regional Science Vol. 81 (no. 2): pages 157-175.

Baeninger, R. (2000). “Região, metrópole e interior: espaços ganhadores e espaços perdedores nas migrações recentes - Brasil.” Textos NEPO. Brazil.

Baldwin, R. E., R. Forslid, P. Martin, G. Ottaviano and F. Robert-Nicoud (2003) “Economic Geography and Public Policy” Princeton University Press, Princeton, N.J., USA.

Banerjee, A. V. and E. Duflo (2005). “Growth Theory Through the Lens of Development Economics.” Economics Handbook of Economic Growth. Vol. 1, Part A: pages 473-552.

Basu, B. (2000). "Rural-Urban Migration, Urban Unemployment and the Structural Transformation of a Dual Economy." Journal of International Trade Vol. 9 (no. 2): pages 137-149.

Bhattacharrya, P. C. (2002). "Rural-Urban Migration in LCD's: A Test of Two Rival Models." Journal of international Development Vol. 14 (no. 7): pages 951-972.

Becker, G. S. (1962). "Investment in Human Capital: A Theoretical Analysis." The Journal of Political Economy Vol. 70 (no. 5): pages 9-49.

Berger, M.C., G. C. Blomquist and J. P. Hoehn (1987). "A Hedonic Model of Interregional Wages, Rents, and Amenity Values." Journal of Regional Science Vol. 27(no. 4): pages 605-620.

Blackaby, D. H. and P. D. Murphy (1991). "Industry Characteristics and Inter-regional Wage Differences." Scottish Journal of Political Economy Vol. 38 (no. 2): pages 142-161.

Blanchflower, D. G. and A. J. Oswald (1995). "An Introduction of the Wage Curve." The Journal of Economic Perspectivies Vol. 9 (no. 3): pages 153-167.

BNDES (2002). “BNDES, um Banco de Idéias - 50 anos refletindo o Brasil.” Rio de Janeiro, BNDES. Brazil.

Brakman, S., H. Garretsen and M. Schramm (2004). "The Spatial Distribution of Wages: Estimating the Helpmen-Hanson Model for Germany." Journal of Regional Science Vol. 44 (no. 3): pages 437-466.

Brulhart, M. and P. Koenig (2006). "New economic geography meets Comecon - Regional wages and industry location in central Europe." Economics of Transition Vol. 14 (no. 2): pages 245-267.

Bronzini, R. and G. d. Blasio (2006). "Evaluating the impact of investments incentives: The case of the Italian Law 488/1992." Journal of Urban Economics Vol. 60 (no. 2): pages 327-349.

Capanema, L. X. d. L. (2006). “A Indústria Farmacêutica Brasileira e a Atuação do BNDES.” BNDES Setorial. N. 23. Rio de Janeiro, BNDES: 193-216.

Carvalho, A. X. Y., D. R. C. Cerqueira and W.J.A. Lobão (2005). “Socioeconomic Structure, Self-Fulfilment, Homicides and Spatial Dependence in Brazil.” IPEA Discussion Paper 1105. Brasília. Brazil.

Carvalho, A. X. Y., D. R. C. Cerqueira, W.J.A. Lobão and R.I. Rodrigues (2007). “Analise dos Custos e Conseqüências da Violência no Brasil.” IPEA Discussion Paper 1284. Brasília. Brazil.

Carvalho, P., C. Salm and J. Saboia (1997). “Produtividade na Indústria Brasileira: questões metodológicas e novas evidencias empíricas.” Pesquisa e Planejamento Economico, Vol. 27 (no. 2): pages 377-396.

Chaudhuri, S. (2000). "Rural-Urban Migration, the Informal Sector, Urban Unemployment, and Development Policies: A Theoretical Analysis." Review of Development Economics Vol. 4 (no. 3): pages 353-364.

Clark, D. E. and J. C. Cosgrove (1991). "Amenities versus labour market opportunities: choosing the optimal distance to move." Journal of Regional Science Vol. 31 (no. 3): pages 311-328.

Combes, P. P., G. Duranton and H.G. Overman (2005). "Agglomeration and the adjustment of the spatial economy." Papers in Regional Science Vol. 83 (no. 3): pages 311-349.

Christaller, W. (1933). “Central Places in Southern Germany.” Jena, Germany: Fischer (English translation by C. W. Baskin, London: Prentice Hall, 1966).

Criscuolo, C., R. Martin, H.G. Overman and J.V. Reenen (2007). “The effect of industrial policy on corporate performance: Evidence from panel data.” Mimeo. London. UK.

Crozet, M. (2004). "Do migrants follow market potentials? An estimation of a new economic geography model." Journal of Economic Geography Vol. 4(no. 4): pages 439-458.

Cullen, J.B. and S. D. Levitt (1999). “Crime, Urban Flight, and the Consequences for Cities.” The Review of Economics and Statistics Vol. 81 (no. 2): pages 159-169.

Cushing, B. J. (1989). "Use and misuse of the allocation rate in models of population migration." The Annals of Regional Science Vol. 23 (no. 1): pages 51-58.

Da Matta, D., U. Deichmann, J.V. Henderson, S.V. Lall and H.G. Wang (2007) "Determinants of city growth in Brazil." Journal of Urban Economics Vol. 62 (no. 2): pages 252-272.

EBRD (2006). “Evaluation Policy of the EBRD. Evaluation-Department.” European Bank for Reconstruction and Development. London. UK

Elizondo, R. L. and P. Krugman (1995). "Trade policy and the Third World metropolis." Journal of Development Economics Vol. 49 (no. 1): pages 137-150.

Fally, T., R. Paillacar and C. Terra (2008). “Economic Geography and Wages in Brazil: Evidence from Micro-Data.” Theme Working Paper, Nº 2008-23. University of Cergy Pontoise. France.

Feijo', C. and P. Carvalho (1994). “Sete Teses Equivocadas sobre o Aumento de Produtividade Industrial nos anos Recentes.” Boletim de Conjuntura, Instituto de Economia, UFRJ. Rio de Janeiro. Brazil.

Ferreira, P. C. and J. L. Rossi (2003). “New Evidence from Brazil on Trade Liberalization and Productivity Growth.” International Economic Review Vol. 44 (no. 4): pages 1383-1405.

Fujita, M., P. Krugman and A.J. Venables (1999) “The Spatial Economy: Cities, Regions, and International Trade.” The MIT Press, Cambridge, Massachusetts, USA.

Gibbons, S. (2004). "The Costs of Urban Property Crime." Economic Journal Vol. 114 (no. 499): pages 441-463.

Glaeser, E. L. and B. Sacerdote (1999). "Why is There More Crime in Cities?" Journal of Political Economy Vol. 107 (no. 6 part. 2): pages S225-S258.

Glaeser, E. L., B. Sacerdote and J.A. Scheinkman (1996). "Crime and Social Interactions." The Quarterly Journal of Economics Vol. 111 (no. 2): pages 507-548.

Goebel, D. (2002). “A Competitividade Externa e a Logística Domestica.” in O Desafio das Exportações. orgarnized by A. C. Pinheiro, R. Markwald and L. V. Pereira. Rio de Janeiro, BNDES: pages 283-366. Brazil.

Greenwood, M. J. (1997). “Internal Migration in Developed Countries.” in Handbook of Population and Family Economics. orgarnized by M. R. Rosenzweig and O. Stark. Amsterdam, New York, Elsevier. 1B: 647-720. USA.

Grilliches, Z., T. J. Klette and J. Moen, et al. (2000). "Do subsidies to commercial R&D reduce market failures? Microeconometric evaluation studies." Research Policy Vol. 29 (no. 4&5): pages 471-495.

Hanson, G. H. (1996). "Localization Economies, Vertical Organization, and Trade." The American Economic Review Vol. 86 (no. 5): pages 1266-1278.

Hanson, G. H. (1997). "Increasing Returns, Trade and the Regional Structure of Wages." The Economic Journal Vol. 107 (no. 440): pages 113-133.

Hanson, G. H. (2005). "Market Potential, Increasing Returns and Geographic Concentration." Journal of International Economics Vol. 67 (no. 1): pages: 1-24.

Hay, D. (2001). "The Post 1990 Brazilian Trade Liberalization and the Performance of Large Manufacturing Firms: Productivity, Market Share and Profits." The Economic Journal Vol. 111 (no. 473): pages 620-641.

Head, K. and T. Mayer (2006). "Regional Wage and Employment Responses to Market Potential in the EU." Regional Science And Urban Economics Vol. 36 (no. 5): pages 573-594.

Heckman, J. J., R. J. LaLonde and J.A. Smith (1999). “The Economics and Econometrics of Active Labour Market Programs.” in Handbook of Labour Economics. organized by O. Ashenfelter, D. E. Card and D. Card, Elsevier. Vol. 3: pages 1865-2097.

Henderson, J. V. (1996). "Ways to Think about Urban Concentration: Neoclassical Urban Systems versus the New Economic Geography." International Regional Science Review , Vol. 19 (no. 1&2): pages 31-36.

Hicks, J. R. (1932). Theory of Wages. London, Macmillan. UK

IANSA “Gun Violence” The Global Crisis. Available at http://www.iansa.org/un/documents/GlobalCrisis07.pdf . Accessed on July 3rd, 2008. London. UK.

Jannuzzi, P. d. M. and N. Jannuzzi (2002). “Crescimento urbano, saldos migratórios e atratividade residencial dos distritos da cidade de São Paulo: 1980-2000.” XIII Encontro Nacional de Estudos Populacionais, Ouro Preto (MG), Associação Brasileira de Estudos Populacionais. Brazil.

Krugman, P. (1980). "Scale Economies, product differentiation and the pattern of trade." American Economic Review Vol. 70 (no. 5): pages 950-959.

Krugman, P. (1991). "Increasing Returns and Economic Geography." The Journal of Political Economy Vol. 99 (no. 3): pages 483-499.

Krugman, P. (1992). “The age of diminishing expectations. U.S. economic policy in the 1990s.” Cambridge MA, MIT Press. USA.

Krugman, P. and A. Venables (1995). "Globalization and the Inequality of Nations." Quarterly Journal of Economics Vol. 60 (no. 4): pages 857-880.

Lacerda, S. M. (2004). "Navegação e Portos no Transporte de Conteineres." Revista do BNDES Vol. 11 (no. 22): pages 215-243.

Lemos, M. B., F.D. Negri and J. A. D. Negri (2006). “Impactos do ADTEN e do FNDCT sobre o desempenho e os esforços tecnológicos das firmas industriais brasileiras” mimeo, IPEA, Brasília. Brazil.

Losch, A. (1940). “The Economics of Location.” Jena, Germany: Fisher (English translation, New Haven, CT: Yale University Press, 1954).

Lucas, R. E. A. (1997). “Internal Migration in Developing Countries.” in Handbook of Population and Family Economics. organized by M. R. Rosenzweig and O. Stark. Amsterdam, New York, Elsevier. pages 721-798. USA.

Martine, G. (1994). “A redistribuição espacial da população brasileira durante a década de 80.” IPEA Discussion Paper. Rio de Janeiro. Brazil.

Melitz, M. (2003) “The impact of trade on intra-industry reallocations and aggregate industry productivity.” Econometrica Vol. 71 (no. 6): pages 1695-1725.

Moteiro Filha, D. C. (1994). “A aplicação de fundos compulsórios pelo BNDES na formação da estrutura setorial da indústria brasileira: 1952 a 1989.” Instituto de Economia Industrial. Doctorate Thesis. Rio de Janeiro, UFRJ. Brazil.

Muendler, M. (2004). "Trade, Technology, and Productivity: A Study of Brazilian Manufacturers 1986-1998." CESifo Working Paper (no. 1148). Munich. Germany.

Murata, Y. (2003). "Product Diversity, Taste Heterogeneity, and Geographic Distribution of Economic Activities: Markets vs. Non-market Interactions." Journal of Urban Economics Vol. 53 (no. 1): pages 126-144.

Murphy, K. M., A. Shleifer and R.W. Vishny (1993). "Why Is Rent-Seeking So Costly to Growth?" The American Economic Review Vol. 83 (no. 2): pages 409-414.

Nassif, A. L. G. (2005). "As Fontes de Mudança da Eficiência Técnica da Indústria Brasileira Apos a Liberalização Comercial: Uma Estimação Econometrica por Dados de Plantas." Revista de Economia Aplicada Vol. 9 (no. 2): pages 301-324.

Nelson, J. M. (1976). “Sojourners versus urbanites: causes and consequences of temporary versus permanent cityward migration in developing countries.” Economic Development and Cultural Change. Vol. 24 (no. 4): pages 721-757.

Nunes, R. and R. Matos (2005). “Migrações internas no Brasil: comparações entre os dados da RAIS e os do Censo Demográfico.” 4° Encontro Nacional Sobre Migração, Associação Brasileira de Estudos Populacionais. Rio de Janeiro, Brazil.

Nuttall, C., D. Eversley and I. Rudder (2002). “The Barbados Crime Survey 2002.” International Comparisons, Barbados Statistical Department. Barbados.

Overman, H. G. and L. A. Winters (2006). "Trade Shocks and Industrial Location: The Impact of EEC accession on the UK." CEP Discussion Paper (no. 558). London. UK.

Pereira, R. O. (2007). "Ação do BNDES sobre o Emprego Formal: Efeito nas Empresas Financiadas." Revista do BNDES Vol. 14 (no. 27): pages 27-42.

Plane, D. A. (2001). "Streams, flows, currents, cycles, and regimes: Propositions on the spatiotemporal nature of migration systems." International Journal of Environmental Creation. Vol. 4: pages 60-73.

Poncet, S. (2006). "Provincial migration dynamics in China: Borders, costs and economic motivations." Regional Science and Urban Economics. Vol. 36 (no. 3): pages 385-398.

Redding, S. and A. J. Venables (2004). "Economic Geography and International Inequality." Journal of International Economics Vol. 62 (no. 1): pages 53-82.

Reiff, L. O. d. A., L.H.R. Rocha and G. A. G. d. Santos (2007). "Emprego Formal, Qualidade de Vida e o Papel do BNDES." Revista do BNDES Vol. 14 (no. 27): pages 5-26.

Reiss, A. J. (1980). “Understanding Changes in Crime Rates.” in Indicators of Crime and Criminal Justice: Quantitative Studies. S. Feinberg and A. J. Reiss. Washington, DC, Bureau of Justice Statistics. USA.

Roback, J. (1982). "Wages, rents and the quality of life." Journal of Political Economy Vol. 90 (no. 6): pages 1257-1278.

Rodriguez-Pose, A. and G. Arbix (2001). "Strategies of waste: bidding wars in the Brazilian automobile sector." International Journal of Urban and Regional Research Vol. 25 (no. 1): pages 134-154.

Rosen, S. (1979). “Wage-based indexes of urban quality of life.” in Current issues in urban economics. organized by P. Mieszkowski and M. Straszheim. Baltimore, Johns Hopkins University Press: 74-104. USA.

Sah, R. K. (1991). "Social Osmosis and Patterns of Crime." Journal of Political Economy Vol. 99 (no. 6): pages 1272-1295.

Santos Silva, J. M. C. and S. Tenreyro (2006). "The log of Gravity." The Review of Economics and Statistics Vol. 88 (no. 4): pages 641-658.

Saracoglu, S. and T. L. Roe (2004). “Rural-Urban Migration and Economic Growth in Developing Countries.” Society for Economic Dynamics – 2004 Meeting Papers. (no. 241). New York. USA.

Schor, A. (2004). "Heterogeneous Productivity Response to Tariff Reduction: Evidence from Manufacturing Firms." Journal of Development Economics Vol. 75 (no. 2): pages 373-396.

Skeldon, R. (1986). “On migration patterns in India during the 1970's.” Population and Development Review. Vol. 12 (no. 4): pages 759-779.

Soares, R. R. (2006). "The Welfare Cost of Violence across Countries." Journal of Health Economics Vol. 25 (no. 5): pages 821-846.

Sousa, F. L. (2002). "A Localização da Indústria Brasileira nas Ultimas Três Décadas." Texto de Discussão do BNDES (no. 95). Rio de Janeiro. Brazil.

Sousa, F. L. d. (2003). "O Papel do BNDES na Distribuição Geográfica da Indústria de Transformação" Revista do BNDES Vol. 10 (no. 19): pages 3-20.

Staiger D. and Stock J. H. (1997). "Instrumental Variables Regression with Weak Instruments" Econometrica Vol. 65 (no. 3): pages 557-586.

Sudekum, J. (2008). "Regional Costs-of-Living with Congestion and Amenity Differences - An Economic Geography Perspective." Annals of Regional Science, forthcoming.

Tabuchi, T. and J.-F. Thisse (2002). "Taste heterogeneity, labour mobility and economic geography." Journal of Development Economics Vol. 69 (no. 1): pages 155-177.
Terra, M. C. (2003). “Credit constraints in Brazilian firms: evidence from panel data.” Revista Brasileira de Economia Vol. 57 (no. 2): pages: 443-464.

Torres Filho, E. T. and F. P. Puga (2006). “Empresas apoiadas pelo BNDES geram mais empregos e pagam mais.” Visão do Desenvolvimento. (no. 17). Rio de Janeiro, BNDES. Brazil.

United Nations (2007). “World Population Prospects: The 2006 Revision and World Urbanization Prospects: The 2007 Revision.” Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat. New York, USA.

UN-Habitat (2007). “Habitat Debate - A safe city is a just city.” in UN - Habitat: 24. Vol. 13. (no. 3). Nairobi, Keynia.

Vasconcellos, I. R. P. and J. I. R. Rangel (2005). “Migrações entre os municípios brasileiros, a partir das informações dos censos demográficos de 1991 e 2000.” 4° Encontro Nacional Sobre Migração, Associação Brasileira de Estudos Populacionais. Rio de Janeiro. Brazil.

Venables, A. (1996). "Equilibrium Locations of Vertically Linked Industries." International Economic Review Vol. 37(no. 2): pages 341-359.

Volpe, C. (2004). "Do Economic Integration and Fiscal Competition Help to Explain Local Patterns?" ZEI Working Paper B04-15. University of Bonn. Germany.

Von Thünen, J.H. (1826). “Der Isolierte Staat in Beziehung auf Landtschaft und Nationalökonomie.” (English translation by C.M.Wartenberg, von Thünen’s Isolated State, Oxford: Pergamon Press 1966).

Waiselfisz, J. J. (2007). “Mapa da Violência dos Municípios Brasileiros.” OEI. Brasilia. Brazil.

Willis, R. J. and C. A. a. R. L. Orley (1986). “Wage determinants: A survey and reinterpretation of human capital earnings functions.” Handbook of Labour Economics, Vol 1: pages 525-602.

Wooldridge, J.M. (2002). “Econometric analysis of cross section and panel data.” MIT Press, Cambridge, Massashutes, USA.

Declaration

I, Filipe Lage de Sousa, confirm that the work presented in this thesis is my own. All chapters in this dissertation have been written entirely on my own. Chapter 4 is a part of a joint work with Gianmarco Ireo Paolo Ottaviano (University of Bologna), but only my empirical contribution is presented in this thesis. No part of this dissertation contains material previously submitted to examiners of any other university. Where information has been derived from other sources, I confirm that this has been duly indicated in the thesis.