NEREUS
Núcleo de Economia Regional e Urbana
da Universidade de São Paulo
The University of São Paulo
Regional and Urban Economics Lab

SHOULD I STAY OR SHOULD I GO? SELECTION ON MIGRATION AND LEARNING IN CITIES IN BRAZIL

Ana Maria Bonomi Barufi
Eduardo Amaral Haddad
Peter Nijkamp

TD Nereus 01-2016
São Paulo
2016

Should I Stay or Should I Go?

Selection on Migration and Learning in Cities in Brazil

Ana Maria Bonomi Barufi, Eduardo Amaral Haddad and Peter Nijkamp

Abstract. The urbanization phenomenon is still going on, especially in developing countries, raising the question of why are cities this successful as productive centers. In this context, agglomeration economies in labor market outcomes have been explored significantly in the past years, even more because longitudinal individual micro-data were made available only recently in many countries. Most studies try to measure static agglomeration externalities, but more recently, there has been an increasing interest in dynamic agglomeration gains (which affect wage growth). Here we consider data for the Brazilian formal labor market from 1995 to 2008 (RAISMIGRA), trying to measure not only static agglomeration externalities, but also dynamic ones. In this sense, we try to understand how migratory movements affect the future wage level of individuals. We calculate the impact of acquired experience in different city sizes over the observed wage, finding that if individuals work more years in bigger cities, they will have higher wages in the future.

1. Introduction

Cities are more important than ever for economic growth and life in society. As discussed before, people living, working, and thinking together in dense areas is a sign of the success of urban agglomerations (Glaeser, 2011). In the United States, Glaeser and Maré (2001) find that in metropolitan areas workers earn 25% more than their non-urban counterparts do. From this perspective, increasing urbanisation is interpreted as a sign of gains from agglomeration (net from costs). In other words, the trade-off between agglomeration economies and congestion costs generates cities (Behrens et al., 2014).

There are three main sources of regional wage disparities in any given moment of time (Combes et al., 2008a): (i) the composition of the local labour market; (ii) the availability of local non-human endowments that can increase productivity; and (iii) agglomeration economies. The third item is related to wage differences that follow from close proximity between firms and workers/consumers, thick labour markets and knowledge spillovers (Duranton and Puga, 2004). Contributing to this discussion, Behrens et al. (2014) divide (i) into two different mechanisms that take place in large cities: a sorting process of more talented individuals, and the selection of the most profitable firms, due to tougher competition. According to them, natural advantages become less important in the modern world.

Most of the literature has been focusing on the static effects of agglomeration economies. In general, the urban wage premium can range from 1% to 11% depending on the sample and the country. The mechanism behind the positive relationship of urban size and wages relies heavily on the assumption that wages, under imperfect competition, are supposed to be higher in places that are more productive (Duranton, 2014a). Thus, agglomerations may generate economic advantages that are not completely internalized by firms and, which may be reflected in higher productivity in larger urban areas.

In addition to their static and instantaneous effect, agglomeration economies can have a long-lasting dynamic impact on productivity (Combes and Gobillon, 2015). Different processes may take place over time: people can learn from each other (De La Roca and Puga, 2014), or they can migrate in search of better opportunities, and return to the original city due to unmatched expectations (De La Roca, 2015; Behrens et al., 2014; Papageorgiou, 2014). These two last processes are confounding factors to the measurement of the dynamic effects of agglomeration economies. They are related to the aforementioned sorting and selection processes and affect the supply of different skills and qualifications in each local labour market over time. On the other hand, learning can influence the way individuals benefit from agglomeration over time.

The empirical literature of agglomeration economies in a developing economy such as Brazil is still very limited, based on estimations at the regional level or with individual cross-sections. As far as the author is aware, there is no study covering dynamic agglomeration economies in the Brazilian context, or for developing countries (Duranton, 2014b). Therefore, this study aims to fill this gap in the empirical research in Brazil, as well as to contribute to the methodological discussion of the literature. This is done by first providing conclusions on the way individuals make their locational choices. Then, a second investigation aims to explore how they benefit over time from agglomeration advantages. Two different models developed elsewhere will provide the basic reduced-form equations to explore these issues.

The first set of results covers an evaluation of the main factors that affect initial and return migration. This is important because it provides a better understanding of the sorting process that may be a confounding factor for the correct measurement of the urban wage premium. Then, the second set of results aims to understand how agglomeration gains vary over time for workers in different-sized cities. A large longitudinal database of administrative reports covering the period from 1995 to 2008 is employed (RAIS-MIGRA), following individuals in their jobs over time, in different firms and cities.

The next sections are organised as follows: Section 2 contains a literature review on static and dynamic externalities, selection in initial and return migration and on the process of learning in cities. Then, Section 3 discusses the theoretical frameworks; Section 4 presents the data; Section 5 provides the main descriptive statistics, Section 6 explores the results; and, finally, Section 7 draws the main conclusions from this analysis.

2. Literature Review

Agglomeration economies may appear in diversified and/or specialised labour markets, with different degrees of interaction between workers and firms. Over time, they can be strengthened by knowledge spillovers, which increase the productive capacity of a certain area. It is in this context that dynamic agglomeration externalities should be understood, as they represent changes over time in local productivity, derived from this increasing capacity for innovation, creating new ways of production, and reducing costs.

In addition to dynamic agglomeration economies, the process of sorting of more productive workers into larger cities can also foster local productivity over time. This means that sorting can be a confounding factor of the advantages derived from agglomerations over time. The next subsections will briefly discuss the main aspects investigated by the literature regarding selection in initial and return migration, and dynamic agglomeration externalities.

2.1. Static *versus* Dynamic Externalities

Static agglomeration economies are defined as once-and-for-all effects of urban size on productivity, with no further impacts in the following periods. Nonetheless, it is also possible that the benefits of agglomeration last longer than one period, with a nonlinear effect on wages and heterogeneous impacts on individuals with different observed and unobserved characteristics. In this context, dynamic agglomeration economies aim to measure the long-lasting effect or permanent impact of urban density on productivity (Combes and Gobillon, 2015). In other words, static advantages relate to the idea that large cities are more productive than smaller ones in a specific moment of time, while dynamic advantages are associated with different trajectories of productivity for each city size over time (Camagni et al., 2015).

The urban economics literature provides strong evidence of the existence of an urban wage premium. This type of analysis can be found, for example, in Combes et al. (2008a), Glaeser and Maré (2001), Melo and Graham (2009), Mion and Naticchioni (2009), Ciccone and Hall (1996), Combes et al. (2010), Groot et al. (2014a), Andersson et al. (2014), and Groot and De Groot (2014b). Literature reviews are provided elsewhere (Puga, 2010, and Rosenthal and Strange, 2004). The elasticity of wages with respect to population density can range from 1% to 11% depending on the sample and the country.

Many studies were recently conducted in different countries because longitudinal individual data on labour market outcomes was made available. However, the literature of agglomeration economies in Brazil is still very restricted, usually based on cross-sectional analysis or aggregated data at the regional level. So far, there has been no study covering dynamic agglomeration economies in the Brazilian context or for an emerging economy, at least according to the author’s knowledge.

Among the possible channels that generate permanent or long-term impacts of agglomeration economies, local productivity growth can be driven by technological spillovers, and individuals may learn more and faster in bigger cities (Lucas, 1988). With regard to the latter aspect, working in bigger cities increases the opportunities of learning (De La Roca, 2014). Glaeser (1999) argues that the rate of interaction between people is accelerated by urban density, leading to an increase in human capital accumulation. This effect is particularly significant for high-skilled people, usually more concentrated in big cities. Yankow (2006) points out that this learning hypothesis requires that individuals stay for a while in the city, meaning that there is no expectation of immediate wage gains when moving to a bigger urban concentration or any sudden wage loss when workers move to smaller cities.

Regarding the empirical estimation of dynamic agglomeration economies, it is still not clear which is the best strategy to analyse and separate them from static effects, due to data limitations. Subsection 2.3 will detail these issues and present the most recent empirical studies.

2.2. Migration and Selection Bias

Migration is very relevant to the spatial redistribution of workers and helps to explain the skill composition of cities over time. Even though migration is expected to equalise real regional wage differentials, it can actually reinforce regional disparities, as migrants are usually more skilled and less risk-averse than stayers are (Greenwood, 1997). For developing countries, special attention has been given to rural–urban movements and to the interaction of migration and formal and informal sector outcomes (Lucas, 1997).

In Brazil, numerous studies have investigated the relationship between internal migration and regional inequality by comparing the wages of migrants and stayers, or of return and permanent migrants. Santos-Junior et al. (2005) suggest that migrants seem to form a positively selected group, with higher earnings in the destination than individuals both at the origin and at the destination. On the other hand, Ramalho and Queiroz (2011) find that return migrants are negatively selected in relation to permanent migrants, but those more qualified obtain wage increases by returning (probably because they find a more suitable occupation for their qualification). Both these studies are based on individual cross-sections.

The proper study of migration should include controls for unobserved and observed individual characteristics, as well as the comparison of individual outcomes over time, and this requires longitudinal databases. In Brazil, the main source for this type of data is RAIS-MIGRA.[1] Based on a sample of individuals from this database, Taveira and Almeida (2014) show that migration flows of qualified migrants between municipalities in Brazil are mostly driven by higher expected income in the destination, population, GDP per capita, the degree of industrialisation and better amenities. Furthermore, a spatial panel technique indicates that the characteristics of neighbouring areas are also relevant in this migration decision. Freguglia et al. (2014) note that after controlling for individual fixed effects, wage differentials between origin and destination states are very relevant in determining the migration of skilled workers. In addition, these migrants look for states with higher prosperity, higher population density, better urban amenities and higher dynamism.

The results found by Freguglia and Procópio (2013) indicate that after controlling for individual unobserved characteristics, changing jobs and interstate migration are the most important factors in determining regional wage differentials. Furthermore, the effect on wages of changing firms is lower than the effect of moving to a different municipality.

However, these studies do not necessarily focus on the relationship between migration and the sorting process that runs parallel to the extraordinary gains derived from large agglomerations. Papageorgeou (2014) and De La Roca (2015) attempt to develop dynamic models of occupational choice that are associated with the different experiences migration can provide to workers. They show that the first decision to migrate differs between skill levels and can be related to the size of the city in both the origin and the destination.

More specifically, Papageorgeou (2014) finds that individuals who have migrated to bigger cities in the United States are more likely to change jobs with a higher frequency, but over time this difference disappears. Furthermore, workers in big cities have a lower propensity to migrate. The positive effect on wages of moving to another city and switching occupation is higher than the one obtained from changing city and staying in the same 3-digit occupation.

[1] This database is provided by the Ministry of Labour as a derived product from the identified RAIS, and will be described in Section 4, and has limited access due to confidentiality issues. The authors are thankful to Professor Raul Silveira-Neto for sharing the data, with all required steps to guarantee the required secrecy.

According to De La Roca (2015), the sorting of more productive workers into bigger cities in Spain may be related to higher earnings in these areas. Their findings suggest that there is a selection by observable characteristics in initial migration, resulting from the fact that the probability of migration is positively associated with initially more productive workers (with higher education attainment and higher occupational skills before moving). The second round of sorting (return migration) is associated with low realized earnings in the first destination (in the case where the first destination is a big city), while high-skilled workers and individuals with high educational attainment are less likely to return.

Therefore, within the context of agglomeration economies, it is important to investigate not only the decision to migrate, but also who benefits the most from agglomeration, and who manages to stay in the destination or has to return. This is crucial in characterising the composition of local labour markets. Berry and Glaeser (2005) show that there is an increasing clustering of skilled workers in cities with originally high skill levels. In the same vein, Elvery (2010) finds that the skill mix employed by firms in smaller cities is lower than the one observed in larger cities. Therefore, large urban areas seem to concentrate a relatively higher percentage of skilled workers, partially because of this migration process.

In summary, larger cities seem to attract and select more skilled individuals. The cost–benefit analysis of living in a big city, with all the costs involved, may not prove positive for low-skilled individuals or for those who are negatively surprised by the income obtained in the destination. Therefore, not only is city size related to the attraction of more qualified workers, but it can also induce less-skilled individuals to leave large agglomerations.

2.3. Learning in Cities

The last subsection emphasised the importance of investigating selection in initial and return migration. These phenomena are very relevant to understanding the changing skill composition of cities. According to Glaeser and Resseger (2010), agglomeration effects are stronger in cities with higher skills. They speculate that the link between human capital and agglomeration economies appears because workers learn more and acquire more skills in big and high-skilled cities, or because technological change can be faster in these places. These perspectives raise the question of how knowledge externalities evolve, and who benefits the most from these gains over time.

The literature on the static urban wage premium has explored at length the main estimation issues and strategies to identify agglomeration economies. As discussed before, there is a fairly well-developed empirical literature on this matter, even though it is still difficult to give a clear interpretation of the results, due to weak links between estimated specifications and theoretical models (Combes and Gobillon, 2015). On the other hand, only recently dynamic agglomeration economies have started to be investigated, requiring longitudinal individual data and representing a further development from static agglomeration estimates.

De La Roca and Puga (2014) conduct one of the most complete empirical studies in dynamic agglomeration economies so far. Their main goal is to separate the possible reasons why firms would pay higher salaries in bigger cities: (i) static advantages; (ii) sorting of more productive workers; (iii) dynamic advantages (cities facilitate learning and experimentation). Therefore, they estimate a Mincerian wage equation, including individual fixed effects to capture the sorting process and the number of years of experience each individual has of working in different city sizes. They estimate this model using a very large Spanish panel data (the whole dataset comprises the period from 1981 to 2009, and the authors evaluate workers' wages from 2004 to 2009). Two restrictions applied to the model discussed in the previous sections are that the authors consider only a few groups of city sizes (instead of allowing a continuous range of city sizes) and that the individual heterogeneity effect is the same in the dynamic and the static contexts.

Carlsen et al. (2013) adopt a similar strategy. Firstly, they focus on the size of the static urban wage premium across education groups, finding increasing elasticities according to educational attainment. However, when unobserved fixed effects are controlled for, these differences disappear, and the common agglomeration elasticity for static effects is around 0.03 (half of the elasticity obtained without controlling for unobserved characteristics). Their data for Norway also allow the estimation of dynamic agglomeration economies, following the strategy of De La Roca and Puga (2014). They separate the effect of experience obtained in large cities from experience obtained elsewhere, and calculate different coefficients for the years worked in the top-ten high wage sectors in comparison to other sectors. Their conclusions are that the aggregated wage premium is not affected by the inclusion of worker experience history. However, the medium-term premium for highly educated workers in high wage sectors is positively affected.

Using a different approach, D'Costa and Overman (2014) consider a large panel (1998–2008) of British workers, aiming to evaluate whether the sorting of high-ability workers can explain urban wage premiums and whether this wage premium is immediately given to workers or it is payed over time through faster wage growth. Their conclusion is that both learning and sorting matter in understanding the effects of cities on wage growth. The main restriction imposed by these authors is that even if they estimate the specification with first differences, allowing for a distinct heterogeneous effect in the static and the dynamic cases, they have to exclude movers from the analysis (therefore, this study does not measure between-city dynamic effects).

A few years earlier, Wheeler (2006) estimated the impact of density on annual wage growth, calculating the within-job and between-job components of this growth, with data for young male workers in the United States. The author's main finding is that there is a positive relationship between wage growth and city size. However, when they control for individual fixed effects, there is no evidence of this urban premium on wage growth.

Dynamic agglomeration gains can be perceived as a reflection of occupational progression. Gordon (2015) shows that the high living costs in large metropolitan areas can only be met by more ambitious individuals. Furthermore, occupational advancement usually results from a combination of the exposure to learning opportunities and the capacity to profit from them (experience and dynamic human capital, respectively). In this context, ambition is supposed to be a key factor, and among the migrants' group, the share of more ambitious individuals seems to be higher. These individuals will seek and enjoy learning and escalating opportunities, which are more likely to appear in the diversified environment of a large urban centre.

The strategy followed by Rattso and Stokke (2011) is slightly different. They look for regional income divergence that could have been caused by migratory movements from the periphery to cities over more than three decades in Norway. The comparison of regional income distribution over time indicates that what actually happened was a process of convergence, contradicting the hypothesis of agglomeration economies generated by migration.

Finally, Yankow (2006) regresses the wage growth against changes in the location of work, separating the mobility effect from the growth effect. The author's conclusions regarding male young workers in the United States is that when they move into cities, they experience a significant wage growth (6 percentage points) in comparison to individuals that stayed in a rural area (with a symmetric effect for out-of-the-city migrants). The main drawback in this author's strategy is that it is not possible to control for sorting on unobservable variables, a limitation that D'Costa and Overman (2014) try to overcome.

In summary, there are different possibilities to estimate dynamic gains from agglomeration economies, but none of them covers all the relevant issues simultaneously. On the one hand, it is possible to identify this effect on the level of wages with movers (De La Roca and Puga, 2014) or with stayers (D'Costa and Overman, 2014), over wage growth (Yankow, 2006; Wheeler, 2006), controlling or not for individual unobserved characteristics (sorting). Here, different strategies will be compared in order to provide as complete an analysis as possible. It is clear, however, that this literature still has to overcome some important limitations.

3. Theoretical Frameworks and Empirical Strategies

The literature on agglomeration economies has been mostly focused on partial equilibrium relationships in the labour market. From this perspective, the interactions between workers and firms can be analysed with a focus on the supply or on the demand side of the labour market. Section 2 discussed the main contributions in the literature to the analysis of selection in migration and of dynamic agglomeration advantages. Now, this section will briefly describe the models that will inspire the empirical analysis of these two subjects in the Brazilian formal labour market.

Here, the sorting process will be studied through the analysis of the migration decision and the importance of individual self-selection in this process. In this context, individuals decide to supply their work in different-sized cities. Such an analysis will provide elements to understand which type of worker is attracted to larger agglomerations. Next, these same workers are faced with the decision of whether or not to stay in the destination. Return migration becomes an additional element to define the sorting of workers. It may determine their resilient behaviour when facing adversities and higher life costs in these larger urban areas, as well as identify who benefits the most from bigger agglomerations.

After that, from the perspective of the demand side of the labour market, wages are evaluated as a reflection of the productivity achieved by firms and how much they are willing to pay to workers of a certain type. Embodying the possibility of learning in bigger cities, these wages may follow a path that will depend on the qualification of the worker, their previous experience in different city sizes and their own unobserved characteristics.

3.1. Selection in Initial and Return Migration

The conceptual framework presented here presupposes that there is a pool of heterogeneous workers initially located in low-density cities L (see De La Roca, 2015). Workers differ only by their skill level s_i. Each worker rents a house of standard quality and spends the remaining income on a numeraire good. In L, housing costs are normalized to zero and the utility of individual i is:

$$U_i{}^L = s_i \tag{1}$$

When workers move to a city of type H (high density), they pay R for rent, receive a skill shock of $\delta_i \sim U[0,2\delta]$ and have a productivity gain of α from working there.

$$U_i{}^H = \alpha(s_i + \delta_i) - R \tag{2}$$

Given that there is a moving cost C, worker i migrates from L to H if and only if $\alpha(s_i + \delta_i) - R - C > s_i$. Thus, workers with low skills will stay in L in equilibrium if:

$$s_i \leq \frac{R + C - \alpha\delta_i}{\alpha - 1} \tag{3}$$

On the other hand, initial migration from H to L will be pursued by workers with skill levels respecting:

$$s_i \leq \frac{R - C - \alpha\delta_i}{\alpha - 1} \tag{4}$$

Then, De La Roca (2015) computes the aggregated effects given the uncertainty of δ_i and the possibility of returning to the initial location by paying an additional moving cost C. The main predictions of the model regarding the selection in initial and return migration from L to H are the following:

- Workers with low s_i do not migrate from L to H;
- Workers with intermediate s_i may migrate from L to H, and depending on the realisation of δ_i, they stay in H or return to L in a second period;
- Workers with high s_i migrate to H and do not return.

Furthermore, workers that are initially in H will observe their realisation of δ_i and decide whether or not to move to L. Then, selection in initial migration from H to L will respect the following conditions:

- Workers with low s_i migrate from H to L and do not return;
- Workers with intermediate s_i migrate from H to L only if they get a bad outcome of δ_i in H;
- Workers with high s_i do not migrate to L.

This simple framework provides the main expected migration movements between low-density (L) and high-density cities (H). With these elements, it is possible to investigate the main factors that affect workers' decision to supply their labour in cities of different sizes.

Therefore, following De La Roca (2015), it is possible to specify a single-exit discrete duration model, defined for the population that may migrate for the first time in each period. Whenever an individual migrates, he or she will be dropped from the population of interest. Each unit of analysis will be considered over time, with a different number of observed characteristics of the individual himself and the place he or she lives. The model to be estimated is based on the probability of migrating at time t, conditional on the fact that the individual has not migrated before t:

$$h(t) = P[T = t|T \geq t] = F[\beta_0(t) + \beta'_1 x(t)] \qquad (5)$$

In this specification, T is the year of migration, F is a logistic cumulative probability function, $x(t)$ is the vector of observed characteristics, β'_1 is a vector of parameters and $\beta_0(t)$ is a duration parameter. The log-likelihood function of a logit model aggregates over time the exit probabilities in each t, with a migration indicator $Y_t = \mathbf{1}(T_i \geq t \geq e_i)$ that equals 0 in every year except the one prior to migration, in which it equals 1, and it is given by:

$$L(\beta) = \sum_{t=1}^{T_i} \left\{ \sum_{i=1}^{N} \mathbf{1}(T_i \geq t \geq e_i) \left[m_i Y_{ti} \log h_i(t) + (1 - m_i Y_{ti}) \log\big(1 - h_i(t)\big) \right] \right\} \qquad (6)$$

where m_i equals 1 when a migration movement is observed and 0 otherwise, e_i is the year of entry in the sample (which will be the same for the whole sample here), and T_i is the number of years until first migration happens. Then, $\hat{\beta}$ will be the maximum likelihood estimator of $L(\beta)$.

Apart from the direct estimation of the first migration decision, this strategy will be the basis for the analysis of the return migration. In this case, the sample will be comprised of individuals who have already migrated once.

3.2. Dynamic Advantages from Agglomeration Economies

The framework adopted by the urban economics literature is responsible for generating the basic productivity specifications that later resulted in the reduced form models estimated there (Combes et al., 2008b). It is possible to depart from that model to obtain dynamic effects in the following way.

A first equation discussed in the literature (Combes et al., 2008a) can be rewritten to encompass the association of individual wages with a composite effect of local productivity, $B_{r,t}$ and the skill level of individuals, $s_{i,t}$. The components of $B_{r,t}$ are the following: pecuniary externalities that appear through local markets (p_j and r_j, referring to agglomeration and dispersion forces respectively), and pure local externalities (that do not appear through a market mechanism and include A_j, a measure of technological externalities).[2]

$$w_{i,t} = B_{r,t} s_{i,t} \tag{7}$$

The inclusion of skills is essential to capture productivity, which is supposed to be directly associated to wages. Although a component of local labour skills could represent the fact that high skills are overrepresented in large cities, it would capture part of the agglomeration effect. Therefore, individual skills seem to be more appropriate, as they are not determined by agglomeration economies. In addition, due to the great difficulty of estimating the effects embedded in $B_{r,t}$ separately, most studies in the literature quantify the overall effect of local productivity on wages (Combes et al., 2011). Furthermore, the relationship between local characteristics and nominal wages is enough to capture the net effect of agglomeration on productivity.

In addition to static effects, in which agglomeration economies have a once-and-for-all impact on productivity, they may also have a long-lasting or even permanent effect. One of the possibilities to evaluate this hypothesis is based on the inclusion of the previous

[2] Scitovsky (1954).

experience of workers in the same region or in other places. This term can account for all the elements included in $B_{r,t-k}$, with $k > 0$, and internalised by the worker.

Departing from Equation 7, Combes and Gobillon (2015) show how dynamic agglomeration effects can be included in the model. Following De La Roca and Puga (2014), it is possible to assume that the log hourly wage of worker i in region r at time t depends on:

$$w_{i,r,t} = \theta_r + \mu_i + \sum_{j=1}^{R} \delta_{j,r} exper_{i,j,t} + X'_{i,t}\boldsymbol{\beta} + \varepsilon_{i,r,t} \tag{8}$$

where θ_r is a region fixed effect (which later on will be related to a measure of agglomeration[3]), μ_i is a worker fixed effect, $exper_{i,j,t}$ represents the experience acquired in region r up until time t, $X_{i,t}$ is a vector of time-varying individual and job characteristics (which include a skill measure), $\boldsymbol{\beta}$ is a vector of parameters and $\varepsilon_{i,r,t}$ is the error term. An initial estimation static agglomeration effect could be based on the following:

$$w_{i,r,t} = \theta_r + X'_{i,t}\boldsymbol{\beta} + \eta_{i,r,t} \tag{9}$$

where individual fixed effects and the term related to previous experience are omitted, and $\eta_{i,r,t}$ is the error term. A pooled OLS estimation of the static urban wage premium embedded in $\theta_{r,t}$ in this restricted model is likely to be biased upwards if individuals with high unobserved ability sort themselves into bigger cities, or if individuals with more valuable previous experience work in bigger cities. From Equation 9, a strategy to address the problem of the sorting of workers is to include individual fixed effects (Combes et al., 2011; Combes et al., 2008a; Combes and Gobillon, 2015; Glaeser and Maré, 2001).

$$w_{i,r,t} = \theta_r + \mu_i + X'_{i,t}\boldsymbol{\beta} + \zeta_{i,r,t} \tag{10}$$

[3] The relationship of density and the spatial wage θ_r will be measured as an average value over the whole period of analysis (following DE LA ROCA; PUGA, 2014). This simplification is necessary because time effects are partially captured by experience (which increases over time at the same speed years pass by.

As mentioned before, the static fixed-effects structure of Equation 10 may still generate an upward-biased estimation of the static urban wage premium, because the term related to experience is still omitted. Therefore, dynamic effects derived from previous experience in the same region or in other places should be included, if the formulation presented in Equation 8 is correct.

When dynamic agglomeration economies are taken into account, Equation 8 will be the basic model to be estimated. However, it is possible to consider variations that will provide a better understanding of who actually profits the most from an additional year of life in each type of city. In this sense, the term related to the experience in region r will interact with a measure of the individual skill level.

In addition to static individual fixed effects, it is possible that unobserved worker characteristics influence wage growth. Thus, Equation 8 is transformed to include an interaction of an individual fixed effect with experience:

$$w_{i,r,t} = \theta_r + \mu_i + \sum_{j=1}^{R} \delta_{j,r} exper_{i,j,t} + \delta_i exper_{i,j,t} + X_{i,t}' \boldsymbol{\beta} + \varepsilon_{i,r,t} \tag{11}$$

In summary, many different alternative models are estimated in this context of dynamic agglomeration economies. Firstly, Equations 9 and 10 provide an initial evaluation of static agglomeration effects, controlling or not for the process of sorting of workers. Then, Equation 11 assesses the importance of dynamic agglomeration effects. These three models generate an estimative of $\hat{\theta}_r$, the region fixed effect that can be defined as a spatial wage (which synthesizes all local characteristics that are relevant for the individual wage). So, this spatial wage is the dependent variable of a second stage,[4] in which a vector of local characteristics G_r, a time dummy, local area and density[5] are the independent variables:

[4] This approach deals with the problem of bias in standard errors discussed by Moulton (1990).

[5] Most empirical studies consider employment instead of population as a measure of the agglomeration because it is a better measure of local economic activity and is easily available for many years (COMBES; GOBILLON, 2015). Ciccone and Hall (1996) defend that density should be included in the place of total employment in this regression, because it deals better with the heterogeneity of the spatial size of the geographical units considered.

$$\hat{\theta}_r = \lambda_t + \varphi \log density_r + \log area_r + G_r\tau + \xi_r \tag{12}$$

This formulation is more similar to the one proposed by D'Costa and Overman (2014). They state that De La Roca and Puga (2014) assume that $\mu_i = \delta_i$ to estimate Equation 11 directly with an iterative process. Even though De La Roca and Puga (2014) follow the two-stage procedure described above that results in the estimation of Equation 12, over here the estimation will be based in only one stage, in order to compare the results directly with the approach described by D'Costa and Overman (2014). Another advantage of this strategy is that it will be possible to measure static and dynamic agglomeration effects at the same level (individual). To do so, Equation 12 is substituted into Equations 8, 9 and 11. In the case of Equation 11, it becomes the following:

$$w_{i,r,t} = \varphi \log density_{r,t} + \log area_r + \mu_i + \sum_{j=1}^{R} \delta_{j,r} exper_{i,j,t} + X'_{i,t}\boldsymbol{\beta} + \varepsilon_{i,r,t} \tag{13}$$

Alternatively, D'Costa and Overman (2014) obtain an estimation of the urban wage premium by evaluating the effect of agglomeration on wage growth. Before presenting their alternative for Equation 11, static effects based on the wage growth are obtained by first-differencing Equation 8:

$$\Delta w_{i,r,t} = \Delta\theta_{r,t} + \sum_{j=1}^{R} \delta_{j,r} exper_{i,j,t} - \sum_{j=1}^{R} \delta_{j,r,t-1} exper_{i,j,t-1} + \Delta X'_{i,t}\boldsymbol{\beta} + \Delta\varepsilon_{i,r,t} \tag{14}$$

When the worker does not move, this equation becomes:

$$\Delta w_{i,r,t} = \delta_{r,r} + \Delta X'_{i,t}\boldsymbol{\beta} + \Delta\varepsilon_{i,r,t} \tag{15}$$

On the other hand, for a worker who moves, the expression becomes much more complicated, mixing static and dynamic effects. Equation 15 provides an estimation of $\delta_{r,r}$ (being in r and staying in r), which represents the value of one additional year spent at r. This approach is not superior to the direct estimation of Equation 8. However, if unobserved workers fixed effects are allowed to affect wage growth, it is possible to estimate them in a way that $\mu_i \neq \delta_i$ (static individual fixed effects are different from dynamic individual fixed effects). In this case, Equation 15 becomes:

$$\Delta w_{i,r,t} = \delta_i + \delta_{r,r} + \Delta X'_{i,t}\boldsymbol{\beta} + \Delta\varepsilon_{i,r,t} \quad (16)$$

It is worth noting that D'Costa and Overman (2014) only exclude from the sample the observations that refer to the exact year the individual has moved. Therefore, by keeping movers, they are still able to identify city effects for them.

Consequently, when there are dynamic agglomeration effects and the individual unobserved heterogeneity can affect the wage growth path, the preferred strategy will be the one adopted by D'Costa and Overman (2014), which is expressed in Equation 16. These results will be compared to Equation 13, a restricted version that does not include the possibility of interaction of individual fixed effects with the wage growth path.

4. Data

The main database considered here is RAIS-MIGRA (Annual Report of Social Information – Migration, from the Ministry of Labour), which consists of identified registration data of all formal firms and their employees in the Brazilian labour market, focusing on the characteristics of the contract. This database provides longitudinal data for all formally employed individuals in the private sector (or part of the public sector, depending on the type of contract), with a significant regional disaggregation (municipal level), from 1995 to 2008. The RAIS-MIGRA is comparable to the identified RAIS[6].

One of the main advantages of this database is that it is a mandatory report, covering the entire formal sector. Because of this, unlike individual self-reporting surveys, there is a smaller risk of wage under-reporting. Nonetheless, the fact that it only covers the formal sector generates a potential drawback in the analysis, as the informal sector in Brazil is very relevant (RAIS represents less than 30% of the workforce in the initial years of the period analysed, reaching around 35% in 2008). Moreover, the reporting process is

[6] The main difference between these two databases is that the process of selecting valid observations is done bythe Ministry of Labour in the case of RAIS-MIGRA, and with steps defined by the author in the case of the identified RAIS.

supposed to be more accurate in the case of bigger firms, which are usually located in larger cities.

The regional unit of analysis is the REGIC area (Area of Influence of Cities),[7] which is a better measure of a labour market area. In fact, this level of analysis is more suitable for the empirical analysis of agglomeration economies, as the estimation is not supposed to be affected by daily commuting (which is very different from moving permanently to another area to look for a job). A few steps were conducted[8] in order to achieve a database representing the dynamics of a competitive labour market, in the manufacturing and service sectors. The total number of observations from 1995 to 2008 before balancing the panel is 43,874,819. The balanced panel, excluding individuals with inconsistent age over time, will be composed of 6,749,778 observations (482,127 individuals in 14 years). Due to limitations of computer processing, a sample of 10%[9] of this database was generated (48,000 individuals observed over 14 years).

One important drawback of the RAIS-MIGRA database is that the classification of occupations is constant over time, and it is not sufficiently disaggregated to be harmonised. Therefore, skills will not be available here, and education attainment will be the only variable capturing the qualification of workers and the complexity of their jobs. In addition, sectors were classified into eight large groups: manufacture; food and accommodation; transport and communication; finance, insurance, pensions and other

[7] REGIC areas (482) aggregate municipalities based on their interaction, mostly associated with daily commuting flows, transportation links and interconnectedness in general terms (IBGE, 2013). From 1995 to 2008 some new municipalities appeared, requiring an adaptation of the original REGIC areas to encompass these new municipalities according to the municipalities they were originated from. The author can provide details of this aggregation process on demand.

[8] There was an initial selection of active contracts in December of each year for male individuals working for private companies in permanent jobs. Then contracts with wages equal to zero, or with less than 20 weekly hours were excluded, as well as contracts with missing information on educational level and occupation. The next step conducted was the selection of individuals aged 18 to 36 years in 1995, with no contradictory age information over time. Finally, only individuals working in the manufacturing and service sectors were kept in the database (sectors 15 to 36, 55 to 74 and 76 to 99, at the 2 digits-CNAE – National Classification of Economic Activities 1.0 classification).

[9] These samples are representative for the following characteristics in the initial year of each database: age group (less than 25 years old, 25 to 29 years old, 30 to 36); density of employment in the REGIC area (less than 1, 1 to less than 2, 2 to less than 5, 5 to less than 10, 10 to less than 20, 20 to less than 50, 50 or more workers per km^2); firm size (up to 4 employees, 5 to 9, 10 to 19, 20 to 49, 50 to 99, 100 to 249, 250 to 499, 500 to 999, 1,000 employees or more); educational level (illiterate, incomplete primary school, complete primary school to incomplete high school, complete high school to incomplete college, college degree or more); and sector (CNAE 1.0 at the 2 digits level).

services; real estate, rents and services to companies; education; health and social services; personal services and other sectors.

Finally, the period of analysis considered for migration models runs from 1995 to 2007 (2008 is omitted as it is just used to define whether the individual has moved to another REGIC area from 2007 to that year). On the other hand, all the estimations regarding static and dynamic agglomeration externalities will be based on the period from 2000 to 2008, and the information from 1995 to 1999 will be used to construct the variable of previous work experience. Table A.1 in the Appendix provides a detailed description of all the variables considered, including methods of calculation and data sources.

5. Descriptive Statistics

This section will present a basic analysis of the main descriptive statistics of the variables that are relevant for the study of the determinants of a migration decision and for the analysis of static and dynamic agglomeration economies.

5.1. Migration Decision

The first aspect to be highlighted here is that this analysis will be limited by the fact that the sample covers only individuals who are working in the formal sector. It is not possible to know what happens to an individual if he or she is not reported in the database in a certain year (they can either be unemployed, out of the labour force, working in the public sector or in the informal sector, or may even be employers). It is also not possible to determine the REGIC area in which the individual is living; only the place he or she is working. Therefore, migration decisions covered here will most likely be the ones in which the individual already has a job offer in another location, and they may represent a selected sample of all migratory movements.

Following the literature[10] and the information available in the database, there is a set of variables that can be considered as explanatory factors for the decision to migrate (initial or return). They are the following: age group, the size of the firm in which the

[10] De La Roca (2015), Papageorgeou (2014), Freguglia et al., (2014).

worker is employed in $t-1$, education attainment,[11] employment density at the origin, sector of employment in $t-1$, macroregion of work in $t-1$, and number of years living in the origin from 1995 onwards. In addition, wage in $t-1$ and the quartile of the annual wage distribution in which the individual was positioned in $t-1$ will be considered as alternatives to education attainment.

As mentioned in the previous section, the REGIC area will be the spatial unit of analysis. In relation to migration, this means that to start working in another REGIC area, the worker has to pay a significant cost. The reason for this is that these areas are supposed to encompass most daily commuting displacements from the municipality of residence to the municipality of work. Therefore, migration is defined by working in a different REGIC area than in the previous year. The main descriptive statistics of the database regarding migration decisions are presented in Table 1.

[11] It would be interesting to include a variable measuring the occupation the individual had in $t-1$, but the database presents a limitation: from 1995 to 2002, occupations were classified according to the CBO 1994 (Código Brasileiro de Ocupações – Brazilian Code of Occupations), and from 2003 to 2008 they were based on the CBO 2002. Furthermore, it is not possible to make them compatible because codes are disaggregated only at the three-digit level until 2002, and at the four-digit level from 2003 onwards.

Table 1. Descriptive Statistics of Individuals According to their Migration Status over the Period 1995–2008

	Non-migrants	1st migration	2nd migration non-return	2nd migration return
ln(hourly wage) - origin	2.64	2.77	3.08	2.74
ln(hourly wage) - destination		2.83	3.14	2.76
Income quantile (origin)				
p25	25.1%	21.0%	19.3%	24.3%
p50	25.0%	21.0%	14.8%	22.1%
p75	25.0%	23.9%	18.5%	23.1%
p100	24.8%	34.0%	47.5%	30.5%
Age	34.2	32.2	34.5	33.6
Tenure in the job (years)	8.4	6.8	5.3	5.8
Education attainment				
Less than 8 years of schooling	34.6%	28.5%	17.3%	29.8%
8 to 10 years of schooling	26.2%	22.9%	17.7%	24.1%
11 to 14 years of schooling	29.5%	32.8%	35.4%	32.7%
15 years of schooling or more	9.7%	15.8%	29.6%	13.5%
Macroregion				
North	1.6%	1.1%	1.8%	1.3%
North-east	9.7%	9.7%	10.2%	10.1%
South-east	65.2%	63.0%	60.7%	67.3%
South	20.3%	22.4%	20.5%	16.3%
Centre-west	3.2%	3.7%	6.8%	5.1%
Sector of activity				
Manufacturing	57.9%	49.0%	34.2%	44.4%
Food and acommodation	2.7%	1.8%	1.5%	2.0%
Transport and communication	10.0%	11.0%	12.6%	11.2%
Finance, insurance, pensions and other services	6.1%	13.9%	21.7%	12.3%
Real estate, rents and services to companies	14.2%	18.4%	25.6%	22.1%
Education	2.0%	1.3%	0.9%	0.5%
Health and social services	3.0%	1.2%	0.9%	1.4%
Personal services and other sectors	4.3%	3.3%	2.6%	6.0%
Firm size				
Up to 4 employees	2.8%	2.0%	2.1%	1.8%
5 to 9	4.8%	3.7%	6.5%	2.7%
10 to 19	5.5%	5.7%	9.7%	4.6%
20 to 49	8.7%	10.3%	13.6%	8.0%
50 to 99	8.4%	9.3%	9.4%	8.0%
100 to 249	14.4%	16.4%	17.3%	15.9%
250 to 499	15.1%	15.4%	13.3%	15.1%
500 to 999	14.9%	15.6%	12.7%	16.8%
1,000 or more employees	25.6%	21.7%	15.5%	27.2%
N	611,693	7,741	1,171	3,389

Source: RAIS-MIGRA.

In general terms, first-time migrants receive a higher salary in the origin than non-migrants, are more qualified, more concentrated in the sectors of finance and real estate, and working originally in medium-sized firms. Then, comparing second-time migrants, those who return to their original REGIC area receive a lower salary in the first location (the location after the first movement, 2.74), and end up receiving a lower salary in the destination than second-time migrants that move to a different REGIC area (2.76 against 3.14). Furthermore, the latter are more qualified (almost 30% of individuals with 15 years of schooling or more, while among migrants who return this group represents only 13.5%). Second-time migrants who return usually worked in bigger firms, and were relatively more concentrated in the South-east (as the first place they have moved into). They are also relatively more concentrated in the manufacturing sector.

Table 2. Descriptive Statistics of the Employment Density at the Origin and Destination

	Non-migrants	1st migration	2nd migration non-return	2nd migration return
Density in the origin (employees per km^2)				
Less than 1	1.8%	3.7%	3.3%	2.5%
1 to less than 2	1.6%	2.5%	4.1%	6.0%
2 to less than 5	5.1%	5.7%	8.9%	7.4%
5 to less than 10	8.9%	10.8%	9.6%	9.1%
10 to less than 20	11.7%	9.8%	14.2%	8.9%
20 to less than 50	21.7%	23.1%	20.1%	24.7%
50 or more	49.3%	44.4%	39.9%	41.4%
Density in the destination (employees per km^2)				
Less than 1		3.0%	4.2%	1.8%
1 to less than 2		4.4%	4.3%	1.8%
2 to less than 5		7.5%	8.5%	3.8%
5 to less than 10		9.0%	11.1%	11.0%
10 to less than 20		11.3%	15.0%	10.9%
20 to less than 50		23.8%	22.3%	21.3%
50 or more		41.0%	34.7%	49.4%
Average distance for the migration (in km)		435.7	462.0	456.4
Years before the 1st migration (from 1995 on)		4.29		
Years after the 1st and before the 2nd migration			2.49	0.92
N	611,693	7,741	1,171	3,389

Density is measured in workers/km^2.
Source: RAIS-MIGRA.

In addition, following Table 2, the distribution of workers according to employment density in the origin and in the destination does not change significantly, except for a decrease in the concentration of workers who live in REGIC areas with a density of 50 or more after they migrate for the first time (from 49.3% to 40.4%). In the opposite direction, second-time migration returning to the original place of work increases the concentration of workers in more dense REGIC areas (from 41.5% to 50.4%). On average, first-time migration takes 4.28 years to happen, while second-time migration happens after 2.53 years for workers who move to a second destination and 1.02 years for those who return to the origin. The average distance travelled to migrate does not change much between each group.

5.2. Dynamic Agglomeration Economies

The estimation of dynamic agglomeration economies will be conducted in one stage, and the main descriptive statistics at the individual level are presented in Table 3. In general, average wages and density at the place of work increased in real terms over time. Consequently, wage growth was positive for most of the years analysed. Furthermore, individuals in the sample became more educated over time, almost doubling the percentage of individuals with College degree or more between 2000 and 2008. This sample also concentrates relatively more individuals working in manufacturing, in big firms and in the South-east.

The experience in each city size is calculated in the following way: from 1995 on, every time the individual is observed in a certain city size group, he or she will get one more year of experience in this city group in the following year. As the database is restricted to evaluate individuals between 2000 and 2008, they are supposed to have at the first year a sum of 5 years of experience in different city sizes. The average experience in each city size will grow every year because of the way this variable is calculated.
Furthermore, as the sample is balanced at the individual level, people will get older over time (there are no new individuals entering the database). After some time, the age group comprising 18 to 24 years old loses its share in the sample.

Table 3. Descriptive Statistics of the Main Variables for the Regressions of Dynamic Agglomeration Externalities

	2000	2001	2002	2003	2004	2005	2006	2007	2008
ln(hourly wage)	2.64	2.64	2.63	2.67	2.72	2.75	2.80	2.84	2.87
Wage growth	.	0.00	-0.01	0.04	0.05	0.03	0.05	0.04	0.03
ln(density)	3.90	3.94	3.99	4.01	4.07	4.12	4.18	4.24	4.29
Experience in cities with density < 5 (in years)	0.53	0.62	0.70	0.78	0.85	0.92	0.99	1.05	1.12
Experience in cities with 5 <= density < 20 (in years)	1.05	1.26	1.47	1.67	1.88	2.08	2.28	2.48	2.67
Experience in cities with density >= 20 (in years)	3.43	4.12	4.83	5.55	6.27	7.00	7.73	8.47	9.21
Education attainment									
Incomplete primary school	35.9%	34.1%	32.7%	31.2%	29.9%	28.5%	27.3%	26.2%	25.0%
Comp. primary school - incompl. middle school	27.1%	26.8%	26.2%	25.6%	25.1%	24.2%	23.8%	23.2%	22.8%
Complete middle school - incomplete college	28.7%	29.9%	31.0%	31.3%	32.5%	33.9%	34.7%	35.4%	36.1%
College degree or more	8.3%	9.2%	10.1%	11.9%	12.6%	13.5%	14.2%	15.2%	16.1%
Age group									
18 to 24	3.6%	1.3%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
25 to 29	21.6%	18.4%	14.6%	10.5%	6.7%	3.6%	1.3%	0.0%	0.0%
30 to 34	32.0%	31.0%	29.2%	26.7%	24.4%	21.5%	18.3%	14.5%	10.3%
35 or more	42.7%	49.3%	56.1%	62.7%	69.0%	74.9%	80.4%	85.4%	89.7%
Firm size									
Up to 4 employees	2.7%	2.7%	2.7%	2.7%	2.8%	2.7%	2.7%	2.6%	2.7%
5 to 9	4.7%	4.8%	4.8%	4.9%	4.8%	4.9%	4.8%	4.8%	4.7%
10 to 19	5.5%	5.5%	5.6%	5.6%	5.5%	5.4%	5.4%	5.5%	5.4%
20 to 49	8.8%	8.7%	8.7%	8.7%	8.5%	8.5%	8.4%	8.2%	8.2%
50 to 99	8.5%	8.8%	8.6%	8.6%	8.4%	8.2%	8.2%	8.0%	7.9%
100 to 249	15.0%	14.8%	14.9%	14.5%	14.0%	13.9%	13.8%	13.5%	13.4%
250 to 499	15.3%	14.9%	14.8%	15.1%	15.0%	14.8%	14.7%	14.2%	14.1%
500 to 999	15.0%	15.9%	15.5%	15.2%	14.9%	15.0%	14.6%	14.3%	14.3%
1,000 or more employees	24.6%	24.0%	24.5%	24.7%	26.2%	26.6%	27.4%	28.8%	29.2%
Sector of activity									
Manufacturing	57.3%	57.2%	57.1%	57.1%	57.4%	57.5%	57.9%	58.1%	57.8%
Food and acommodation	2.7%	2.6%	2.6%	2.6%	2.6%	2.6%	2.6%	2.6%	2.6%
Transport and communication	10.1%	10.1%	10.1%	10.0%	10.1%	10.1%	10.2%	10.2%	10.2%
Finance, insurance, pensions and other services	6.3%	6.3%	6.2%	6.2%	6.2%	6.2%	6.1%	6.3%	6.2%
Real estate, rents and services to companies	14.5%	14.6%	14.8%	14.9%	14.5%	14.5%	14.0%	13.9%	14.0%
Education	1.9%	1.9%	1.9%	2.0%	2.0%	2.0%	2.1%	2.0%	2.1%
Health and social services	3.0%	3.0%	2.9%	2.9%	2.9%	3.0%	3.0%	2.9%	3.0%
Personal services and other sectors	4.3%	4.3%	4.3%	4.3%	4.3%	4.2%	4.2%	4.1%	4.1%
Macro-region									
North	1.6%	1.6%	1.6%	1.6%	1.6%	1.6%	1.6%	1.6%	1.6%
North-east	9.7%	9.8%	9.7%	9.7%	9.7%	9.8%	9.8%	9.8%	9.7%
South-east	65.2%	65.2%	65.3%	65.2%	65.2%	65.1%	65.1%	65.0%	65.0%
South	20.2%	20.1%	20.2%	20.3%	20.3%	20.3%	20.3%	20.4%	20.4%
Centre-West	3.3%	3.3%	3.2%	3.2%	3.2%	3.2%	3.2%	3.3%	3.3%
N	48,000	48,000	48,000	48,000	48,000	48,000	48,000	48,000	48,000

Source: RAIS-MIGRA.

6. Results

This section will discuss the main determinants of the migration decision and the effects of static and dynamic agglomeration economies over individual wages.

6.1. Migration

This subsection will discuss the main results found in the analysis of selection in initial and return migration and in the measurement of dynamic agglomeration economies. The first set of results refers to the estimation of the determinants of the first migration. As noted previously, the database is composed only of employed individuals, which means that migrants will necessarily have to find a job at the destination. In addition, the variable of first-time migration is defined as follows: equals 0 if the individual has never moved; equals 1 in the year the individual moves. After this first move, the individual is dropped from the database.

The logit models in column 1 of Table 4 indicate that workers with 15 years of schooling or more are 1.62 times more likely to migrate than workers with less than 8 years of schooling. In addition, a 10% increase in the logarithm of density in the origin reduces the probability of outmigration by 0.5%.[12] Furthermore, individuals in the South-east are 1.83 times more likely to outmigrate from their original REGIC area than workers in the Northern region of the country.

Each additional year of age increases slightly the probability of outmigration (5.3%), while an additional year of tenure at the job decreases this probability by 3.1%. Moreover, workers in the financial and real estate sectors are much more likely to migrate than workers in other sectors.

[12] This value is calculated as $10\% * (0.844 - 1)/e$.

Table 4. Logit Models for the Probability of First-time Migration, 1995 to 2007

(continues)	Model 1 Logit	Model 2 Logit	Model 3 Logit	Model 4 Logit	Model 5 Logit	Model 6 Logit
ln(density in the origin)	0.844***	0.834***	0.826***			
	(0.011)	(0.011)	(0.011)			
Density in the origin (employees per km²) - omitted less than 1						
1 to less than 2				0.605***	0.597***	0.589***
				(0.061)	(0.060)	(0.059)
2 to less than 5				0.369***	0.372***	0.364***
				(0.032)	(0.032)	(0.032)
5 to less than 10				0.430***	0.443***	0.430***
				(0.034)	(0.035)	(0.034)
10 to less than 20				0.325***	0.330***	0.318***
				(0.027)	(0.028)	(0.027)
20 to less than 50				0.358***	0.354***	0.343***
				(0.030)	(0.029)	(0.028)
50 or more				0.307***	0.299***	0.286***
				(0.027)	(0.027)	(0.026)
Education attainment - omitted less than 8 years of schooling						
8 to 10 years of schooling	1.164***			1.170***		
	(0.038)			(0.038)		
11 to 14 years of schooling	1.589***			1.596***		
	(0.050)			(0.050)		
15 years of schooling or more	2.622***			2.603***		
	(0.106)			(0.105)		
ln(hourly wage) - origin			1.523***			1.507***
			(0.026)			(0.025)
Income quantile (origin) - omitted p25						
p50		1.087**			1.085**	
		(0.040)			(0.040)	
p75		1.386***			1.378***	
		(0.052)			(0.052)	
p100		2.219***			2.180***	
		(0.086)			(0.085)	
Age	1.053***	1.037**	1.030*	1.053***	1.038**	1.031*
	(0.019)	(0.018)	(0.018)	(0.019)	(0.018)	(0.018)
Age squared	0.999***	0.999***	0.999**	0.999***	0.999***	0.999**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Tenure in the job	0.969***	0.960***	0.960***	0.970***	0.961***	0.961***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Macroregion - omitted North						
North-east	2.597***	2.840***	3.103***	3.082***	3.305***	3.593***
	(0.317)	(0.348)	(0.381)	(0.389)	(0.419)	(0.456)
South-east	2.832***	2.910***	2.975***	3.306***	3.316***	3.378***
	(0.336)	(0.346)	(0.354)	(0.406)	(0.408)	(0.417)
South	2.703***	2.885***	2.915***	3.539***	3.704***	3.742***
	(0.320)	(0.342)	(0.346)	(0.447)	(0.469)	(0.475)
Centre-west	1.946***	1.929***	1.966***	2.403***	2.333***	2.379***
	(0.249)	(0.247)	(0.252)	(0.319)	(0.310)	(0.317)

(end)	Model 1 Logit	Model 2 Logit	Model 3 Logit	Model 4 Logit	Model 5 Logit	Model 6 Logit
Sector of activity - omitted Manufacturing						
Food and acommodation	1.090	1.240**	1.345***	1.075	1.215**	1.312***
	(0.098)	(0.113)	(0.122)	(0.097)	(0.110)	(0.119)
Transport and communication	1.510***	1.585***	1.564***	1.507***	1.576***	1.554***
	(0.061)	(0.065)	(0.063)	(0.061)	(0.064)	(0.063)
Finance, insurance, pensions and other services	2.946***	3.175***	3.058***	2.879***	3.115***	3.005***
	(0.122)	(0.130)	(0.126)	(0.119)	(0.127)	(0.124)
Real estate, rents and services to companies	2.121***	2.353***	2.364***	2.105***	2.328***	2.335***
	(0.076)	(0.086)	(0.086)	(0.076)	(0.085)	(0.085)
Education	0.702***	0.850	0.847	0.702***	0.847	0.845*
	(0.072)	(0.086)	(0.086)	(0.072)	(0.086)	(0.086)
Health and social services	0.471***	0.566***	0.565***	0.463***	0.557***	0.556***
	(0.049)	(0.059)	(0.059)	(0.048)	(0.058)	(0.058)
Personal services and other sectors	1.004	1.106	1.105	0.987	1.087	1.087
	(0.068)	(0.074)	(0.074)	(0.066)	(0.073)	(0.073)
Firm size - omitted up to 4 employees						
5 to 9	1.048	1.025	1.004	1.048	1.025	1.002
	(0.106)	(0.104)	(0.102)	(0.106)	(0.104)	(0.102)
10 to 19	1.370***	1.363***	1.319***	1.385***	1.382***	1.339***
	(0.132)	(0.131)	(0.127)	(0.133)	(0.133)	(0.129)
20 to 49	1.655***	1.648***	1.575***	1.672***	1.670***	1.597***
	(0.150)	(0.150)	(0.143)	(0.152)	(0.152)	(0.145)
50 to 99	1.824***	1.810***	1.713***	1.840***	1.833***	1.737***
	(0.167)	(0.166)	(0.157)	(0.168)	(0.168)	(0.159)
100 to 249	1.925***	1.911***	1.784***	1.938***	1.931***	1.806***
	(0.170)	(0.169)	(0.158)	(0.171)	(0.171)	(0.160)
250 to 499	1.762***	1.740***	1.618***	1.782***	1.765***	1.645***
	(0.156)	(0.155)	(0.144)	(0.158)	(0.157)	(0.147)
500 to 999	1.852***	1.824***	1.698***	1.855***	1.835***	1.712***
	(0.164)	(0.162)	(0.151)	(0.165)	(0.163)	(0.153)
1,000 or more employees	0	1.470***	1.378***	1.548***	1.478***	1.387***
	(0.136)	(0.130)	(0.122)	(0.136)	(0.130)	(0.123)
Years without migrating (from 1995 on) - omitted 0						
1 year	0.774***	0.795***	0.764***	0.765***	0.783***	0.752***
	(0.032)	(0.032)	(0.031)	(0.031)	(0.032)	(0.031)
2 years	0.569***	0.599***	0.559***	0.557***	0.584***	0.545***
	(0.026)	(0.028)	(0.026)	(0.026)	(0.027)	(0.025)
3 years	0.456***	0.491***	0.458***	0.441***	0.472***	0.441***
	(0.023)	(0.025)	(0.024)	(0.023)	(0.024)	(0.023)
4 years	0.475***	0.527***	0.496***	0.464***	0.511***	0.481***
	(0.025)	(0.028)	(0.026)	(0.025)	(0.027)	(0.026)
5 years	0.410***	0.467***	0.430***	0.411***	0.464***	0.428***
	(0.023)	(0.026)	(0.024)	(0.023)	(0.027)	(0.025)
6 years	0.410***	0.477***	0.441***	0.404***	0.466***	0.431***
	(0.024)	(0.028)	(0.025)	(0.024)	(0.027)	(0.025)
7 years	0.294***	0.353***	0.328***	0.288***	0.342***	0.318***
	(0.019)	(0.023)	(0.022)	(0.019)	(0.023)	(0.021)
8 years	0.286***	0.354***	0.324***	0.279***	0.342***	0.314***
	(0.020)	(0.024)	(0.022)	(0.019)	(0.024)	(0.022)
9 years	0.279***	0.353***	0.318***	0.271***	0.340***	0.306***
	(0.020)	(0.025)	(0.022)	(0.019)	(0.024)	(0.022)
10 years	0.450***	0.584***	0.521***	0.435***	0.559***	0.498***
	(0.028)	(0.037)	(0.033)	(0.028)	(0.036)	(0.032)
11 years	0.419***	0.557***	0.488***	0.404***	0.531***	0.465***
	(0.028)	(0.037)	(0.032)	(0.027)	(0.036)	(0.031)
12 years	0.352***	0.480***	0.414***	0.344***	0.464***	0.401***
	(0.025)	(0.035)	(0.030)	(0.025)	(0.034)	(0.029)
Pseudo R^2	0.0537	0.0530	0.0539	0.0541	0.0532	0.0540
N	564,306	564,306	564,306	564,306	564,306	564,306

Odds ratios (exponentiated coefficients) are reported for each model. Additional controls: constant, percentage of workers in each of the economic sectors in the whole labour market of each REGIC area. The reference category is stayers, and once individuals move they are dropped from the sample. Density is measured as the number of workers per km^2. Standard errors are presented in parentheses. Significance levels: p<0.10, ** p<0.05, *** p<0.01.

Source: Author’s own calculations.

One final result shown in column 1 of Table 4 is that if workers stay in a certain place for longer, the probability that they will outmigrate becomes even lower. Therefore, the chance that a worker who is being observed in a specific REGIC area will migrate after 5 years is 59.0% lower than the chance of a worker who has just arrived deciding to migrate. This finding is also driven by the fact that workers are observed for the first time in 1995 (but they could have been working before that). As time passes by, the remaining individuals in the sample (those who have not migrated yet) will become more and more selected with characteristics of stayers.

In column 2, education attainment is substituted by the position the worker occupied in the distribution of wages of the sample in the previous year. Thus, workers who were in the second quartile had a chance of migration that was 8.7% higher than workers in the lower quartile. Moreover, workers in the upper quartile present a chance of migrating that is 122% higher than workers in the lower quartile. When the individual wage is considered in the place of education attainment (column 3), a very similar result is found – individuals with higher wages will be more likely to outmigrate (a 10% increase in the logarithm of the hourly wage increases the probability of migration in 1.9%). Finally, columns 4 to 6 show that when different density groups are considered (instead of the observed local density), the probability of outmigration decreases for denser cities. The other main results remain unchanged.

Therefore, workers who migrate for the first time seem to be positively selected (higher education attainment, higher initial wage, working in sectors with higher knowledge intensity), but the size of the local labour market decreases the probability of migration, as expected. This means that migration movements in the country are expected to happen from smaller to larger cities, where job opportunities are easier to find and wages are supposedly higher.

The results presented in Table 4 referred to all possible migration movements. However, in Table 5 the estimations are measuring the probability that a worker will move to a REGIC area with a density of at least 10 workers per squared kilometre, with an increase in local density from the previous location. This set of models aims to investigate the main factors that increase the chance that a worker will move to a denser urban area.

Comparing these results, it is noticeable that local density in the origin becomes even more important as a barrier to new migration movements (now a 10% increase in the logarithm of employment density will decrease the probability of migration to a large city in 1.5% - which is compared to a decrease of 0.5% in the previous table). This is explained by the fact that workers who are already in large urban areas have a lower incentive to move to a similar REGIC area in terms of the complexity of the local labour market.

Table 5. Logit Models for the Probability of First-time Migration towards a REGIC Area with a Density Higher than 10 workers/km^2 and with an Increase in Density, 1995 to 2007

	Model 1 Logit	Model 2 Logit	Model 3 Logit	Model 4 Logit	Model 5 Logit	Model 6 Logit
ln(density in the origin)	0.580***	0.570***	0.563***			
	(0.011)	(0.011)	(0.011)			
Density in the origin (employees per km^2) - omitted Less than 1						
1 to less than 2				0.453***	0.437***	0.429***
				(0.057)	(0.055)	(0.054)
2 to less than 5				0.300***	0.300***	0.291***
				(0.032)	(0.032)	(0.031)
5 to less than 10				0.335***	0.342***	0.330***
				(0.033)	(0.034)	(0.033)
10 to less than 20				0.286***	0.284***	0.274***
				(0.029)	(0.029)	(0.028)
20 to less than 50				0.283***	0.270***	0.260***
				(0.029)	(0.028)	(0.027)
50 or more				0.122***	0.115***	0.109***
				(0.014)	(0.014)	(0.013)
Education attainment - omitted Less than 8 years of schooling						
8 to 10 years of schooling	1.233***			1.229***		
	(0.060)			(0.060)		
11 to 14 years of schooling	1.690***			1.677***		
	(0.080)			(0.079)		
15 years of schooling or more	3.514***			3.415***		
	(0.214)			(0.207)		
ln(hourly wage) - origin			1.874***			1.831***
			(0.047)			(0.046)
Income quantile (origin) - omitted p25						
p50		1.312***			1.278***	
		(0.073)			(0.071)	
p75		1.973***			1.903***	
		(0.112)			(0.108)	
p100		3.449***			3.292***	
		(0.201)			(0.193)	
Pseudo R2	0.0906	0.0929	0.0952	0.0819	0.0837	0.0857
N	559,925	559,925	559,925	559,925	559,925	559,925

Odd ratios (exponentiated coefficients) are reported for each model. Additional controls: constant, percentage of workers in each of the economic sectors in the whole labour market of each REGIC area, firm size, macroregion, years before migration, sector of activity, age, age squared and tenure at the job, years before migration since 1995. The reference category is stayers, and once individuals move they are dropped from the sample. Individuals who move at any moment to another REGIC area with a density lower than 10 or who move to a REGIC area with a lower density than the previous place of work are excluded from the sample. Density is measured as the number of workers per km^2. Standard errors are presented in parentheses. Significance levels: $p<0.10$, ** $p<0.05$, *** $p<0.01$.
Source: Author's own calculations.

Furthermore, the average income in the origin becomes an even more important factor: workers in the upper quartile are 2.45 times more likely to migrate to a dense REGIC area than workers in the lower quartile (column 2). The results presented in column 1 indicate that more educated workers are relatively more likely to migrate to large urban areas, especially when compared to all possible initial migration movements. Noteworthy, more educated workers have a higher propensity to migrate to larger cities than to migrate to any city (their coefficients in Table 5 are higher than the ones observed in Table 4).

Tables 4 and 5 provide the conclusion that there is a relevant process of self-selection in first-time migration in the formal labour market in Brazil. This selection becomes even more important when only migration to denser urban areas is taken into account. Therefore, the way incentives are distributed in space makes it more likely that regional inequalities will increase, as cities with a higher percentage of skilled workers will attract more of these high-qualified individuals.

In addition to that, it is important to note that, in general, migrants are more skilled and have higher wages in the origin than stayers. Therefore, even if a worker is moving to a less dense city, he or she is more likely to be more productive than a similar worker who stayed in the original REGIC area. Migration is a relevant instrument of selection of risk-takers, entrepreneurs, young and skilled workers.

After analysing the determinants of first-time migration, it is possible to identify which are the main characteristics of workers who decide to migrate for a second time. These migrants have the option of returning to the original REGIC area (“return”) or of going to another place (“move on”). Consequently, it is necessary to estimate a multinomial logit with three options: staying in the second place of work, moving to a third place, or returning to the original place of work. Table 5 presents a comparison of these events for workers who have already migrated once. Therefore, workers who have moved to a REGIC area during the previous years and decided to stay there afterwards compose the base group.

The main results of Table 6 are the following: workers who decide to return to their original REGIC areas are less educated than those who decide to stay or to move to another REGIC area (having 15 years of schooling or more decreases the probability of returning in 48.3%). Furthermore, wages in the second location (after the first migration) are negatively associated to the probability of moving again, but less so for moves in direction to a third location. Local density is negatively associated to the probability of leaving a certain REGIC area, independently of the direction of this movement. Finally, the larger the distance the worker had to cover in the first migration, the higher the probability he or she will move to another place.

Table 6. Multinomial Logit Models for the Determinants of Second-time Migration (for individuals who have already migrated once), 1996 to 2007

	Model 1		Model 2		Model 3		Model 4	
	Return	**Move on**	**Return**	**Move on**	**Return**	**Move on**	**Return**	**Move on**
ln(density in the origin)	0.744***	0.916**	0.746***	0.928**				
	(0.014)	(0.033)	(0.015)	(0.033)				
Density in the origin (employees per km^2) - omitted less than 1								
1 to less than 2					2.270***	0.908	2.178***	0.898
					(0.364)	(0.216)	(0.353)	(0.214)
2 to less than 5					0.981	0.996	0.956	0.992
					(0.155)	(0.218)	(0.152)	(0.219)
5 to less than 10					0.753*	0.787	0.738*	0.787
					(0.118)	(0.173)	(0.116)	(0.174)
10 to less than 20					0.419***	0.748	0.404***	0.738
					(0.068)	(0.167)	(0.066)	(0.166)
20 to less than 50					0.566***	0.727	0.534***	0.691
					(0.092)	(0.171)	(0.088)	(0.163)
50 or more					0.285***	0.670*	0.280***	0.680
					(0.047)	(0.163)	(0.047)	(0.166)
Education attainment - omitted less than 8 years of schooling								
8 to 10 years of schooling	0.878**	1.159			0.897**	1.167		
	(0.046)	(0.118)			(0.047)	(0.119)		
11 to 14 years of schooling	0.685***	1.357***			0.699***	1.361***		
	(0.035)	(0.126)			(0.036)	(0.126)		
15 years of schooling or more	0.517***	1.941***			0.509***	1.940***		
	(0.034)	(0.197)			(0.034)	(0.196)		
ln(hourly wage)$_t$ at the 2nd location			0.490***	0.870**			0.492***	0.866**
			(0.019)	(0.055)			(0.019)	(0.055)
ln(hourly wage) at the 1st location			1.813***	1.712***			1.793***	1.722***
			(0.072)	(0.111)			(0.071)	(0.111)
ln(distance in the first migration)			1.070***	1.087**			1.059***	1.090**
			(0.022)	(0.038)			(0.022)	(0.038)
Adjusted R^2	0.0527		0.0629		0.0560		0.0658	
N	31,403		31,353		31,403		31,353	

Relative risk ratios (exponentiated coefficients) are reported for each model. Additional controls: constant, percentage of workers in each of the economic sectors in the whole labour market of each REGIC area, firm size, macrorregion, years before migration, sector of activity, age, age squared and tenure at the job, years before migration since 1995. The reference category is first-time migrants who become stayers, and once individuals move for the second time, they are dropped from the sample. Individuals who move at any moment to another REGIC area with density lower than 10 or who move to a REGIC area with lower density than the previous place of work are excluded from the sample. Density is measured as the number of workers per km^2. Standard errors are presented in parentheses. Significance levels: $p<0.10$, ** $p<0.05$, *** $p<0.01$.
Source: Author's own calculations.

Therefore, second migration seems to reinforce the effects of first-time migration. Workers are positively selected to migrate to a third place, but those who decide to return to the original REGIC area are less qualified, with lower salaries in the place they were living after the first migration. A final aspect to be studied regarding selection in initial and return migration refers to the correlation between wages in the origin and migration decisions. It is not possible to discuss causality here, because wages in $t-1$ are correlated to migration movements in t.

Table 7. Partial Correlation of the ln of Hourly Wages in t-1 and Migration Decisions in t, 1996 to 2007

	1st migration		2nd migration	
	Model 1	**Model 2**	**Model 3**	**Model 4**
Migration towards a dense city	0.151***	0.009*		
	(0.010)	(0.005)		
Migration towards a less dense city	0.037***	-0.008*		
	(0.008)	(0.005)		
Migration towards a dense city - return			0.031***	0.007
			(0.016)	(0.011)
Migration towards a less dense city - return			-0.196***	-0.029**
			(0.016)	(0.011)
Migration towards a dense city - no return			0.169***	0.017
			(0.028)	(0.015)
Migration towards a less dense city - no return			0.017***	0.005
			(0.024)	(0.013)
ln(hourly wage)$_{t-2}$		0.869***		0.885***
		(0.001)		(0.003)
Adjusted R^2	0.5361	0.8869	0.5647	0.9106
N	564,306	516,306	31,403	23,999

* Additional controls: constant, percentage of workers in each of the economic sectors in the whole labour market of each REGIC area, firm size, macrorregion, years before migration, sector of activity, age, age squared and tenure at the job, years before migration since 1995. Individuals who move at any moment to another REGIC area are excluded from the sample. Density is measured as the number of workers per km^2. Standard errors are presented in parentheses. Significance levels:* p<0.10, ** p<0.05, *** p<0.01.
Source: Author's own calculations.

Table 7 presents these correlations. In relation to first-time migration (Model 1), workers who have decided to migrate to a dense city (10 or more workers per squared kilometres) used to receive a higher wage in the origin. Even workers who have migrated to a less dense city (less than 10 workers per squared kilometres) received in average a higher wage than stayers. In Model 2, the wage two years before migrating is included, and the coefficients are now supposed to capture the correlation between migration and wage growth. In fact, workers who have decided to migrate to a less dense city had seen a decrease in wages prior to taking this decision, while those who have migrated to a dense city had seen an increase in their wages.

Regarding the second migration, the decision to migrate to a dense city is related to higher wages in the previous period, even if workers are returning to their original REGIC area. Furthermore, in the case workers decide to migrate to a third place, their wages in the previous period were higher, and those who decide to go to a dense city were supposed to have the higher original wages.

Even if these results are not indicating any causality, they seem to show that workers who migrate were originally better off in their first location than stayers, and workers who migrate for the second time to a different location are positively selected once again. Migration in the formal labour market in Brazil seems to have a very important role in the sorting of more skilled individuals in large urban areas.

6.2. Dynamic Agglomeration Economies

The literature review presented in Section 2.2 indicates that there is not a well defined strategy to measure dynamic agglomeration economies. Here, two different possibilities will be tested. The first is based on the estimation of Equation 13, with a special focus on the effects of work experience in cities over the level of wages. On the other hand, the second approach is based on the investigation of the main factors that affect the wage growth (Equation 16).

A first aspect to be highlighted in Table 8 is that the process of sorting seems to be extremely relevant to explain what would be otherwise accredited to the urban wage premium. The elasticity of employment density in relation to individual wages decreased from 0.091 to 0.006 when individual fixed effects were taken into account (Models 1 and 2). This is an indication that a significant share of the positive relationship between wages and local density can actually be explained by the migration of self-selected workers to large urban areas, as discussed in the previous subsection.

Moreover, years of experience in each type of city are included in Model 3. It is possible to find positive and significant coefficients for all of them, with a higher coefficient for the years worked in less dense cities (density < 5). When these years of previous experience are combined with the present place of work (Model 4), individuals working in less dense cities who had previous experience in denser areas will benefit the most from these gains.

In other words, workers that are now in cities with density lower than 5 and have worked in cities with density higher than 5 and lower than 20 are likely to obtain an average wage increase of 1.8% for each additional year of previous experience. In the same way, if previous experience was obtained in areas with density higher than 20, one additional year of experience increases wages in cities with density lower than 5 in 0.8%. In conclusion, previous work experience is more valuable in less dense cities, with a higher effect for the years worked in cities with medium density level.

Table 8. Estimation of the Dynamic and Static City-size Earnings Premia (dependent variable is the logarithm of the hourly wage), 2000 to 2008

	OLS Model 1	FE Model 2	FE Model 3	FE Model 4
ln(density)	0.091***	0.006***	0.005***	0.009***
	(0.001)	(0.001)	(0.001)	(0.002)
ln(area)	0.027***	0.006***	0.006***	0.004*
	(0.001)	(0.002)	(0.002)	(0.002)
Experience in cities with density < 5 (in years)			0.034***	0.039***
			(0.001)	(0.003)
Experience in cities with density < 5 (in years) ^{2}				0.000*
				(0.000)
Experience in cities with 5 <= density < 20 (in years)			0.031***	0.028***
			(0.000)	(0.001)
Experience in cities with 5 <= density < 20 (in years) ^{2}				0.000***
				(0.000)
Experience in cities with density >= 20 (in years)			0.026***	0.029***
			(0.000)	(0.001)
Experience in cities with density >= 20 (in years) ^{2}				0.000***
				(0.000)
Experience in cities with 5 <= density < 20 (in years)*Now in city with density < 5				0.018***
				(0.002)
Experience in cities with 5 <= density < 20 (in years)*Now in city with density >= 20				0.001**
				(0.000)
Experience in cities with density >= 20 (in years)*Now in city with density < 5				0.008***
				(0.002)
Experience in cities with density >= 20 (in years)*Now in city with 5 <= density < 20				0.000
				(0.001)
Education attainment - omitted less than 8 years of schooling				
8 to 10 years of schooling	0.198***			
	(0.002)			
11 to 14 years of schooling	0.595***			
	(0.002)			
15 years of schooling or more	1.313***			
	(0.003)			
Region dummies	Yes	Yes	Yes	Yes
Age group dummies	Yes	No	No	No
Firm size dummies	Yes	Yes	Yes	Yes
Sector dummies	Yes	Yes	Yes	Yes
Worker fixed effects	No	Yes	Yes	Yes
R^2	0.5207			
R^2 within		0.1375	0.1382	0.1386
R^2 between		0.1773	0.1102	0.1474
N	432,000	432,000	432,000	432,000

* All models include a constant term. Density is measured as the number of workers per km^2. Standard errors are presented in parentheses. Significance levels:* p<0.10, ** p<0.05, *** p<0.01.
Source: Author's own calculations.

Table 9. Regression for the Urban Wage Growth Premium (dependent variable is the individual wage growth in one year), 2001 to 2008

	OLS Model 1	OLS Model 2	FE Model 3	FE Model 4	OLS Model 5	OLS Model 6	FE Model 7	FE Model 8
Density >= 5 & density < 20	-0.003*	-0.006***	-0.011*	-0.014**	-0.011***	-0.012***	-0.008	-0.010
	(0.002)	(0.002)	(0.006)	(0.007)	(0.003)	(0.003)	(0.006)	(0.007)
Density >= 20	-0.008***	-0.011***	-0.004	-0.002	-0.010***	-0.012***	-0.002	0.003
	(0.002)	(0.002)	(0.007)	(0.008)	(0.003)	(0.004)	(0.007)	(0.009)
Experience in cities with density < 5 (in years)					-0.027	-0.036	0.001	0.004***
					(0.038)	(0.038)	(0.001)	(0.001)
Experience in cities with 5 <= density < 20 (in years)					-0.026	-0.036	0.004***	0.006***
					(0.038)	(0.038)	(0.001)	(0.001)
Experience in cities with density >= 20 (in years)					-0.027	-0.036	0.004***	0.007***
					(0.038)	(0.038)	(0.000)	(0.000)
Education attainment (incomplete primary school omitted)								
Comp. primary school - incompl. middle school		0.002	-0.006**	-0.010***		0.002	-0.006**	-0.010***
		(0.001)	(0.003)	(0.004)		(0.001)	(0.003)	(0.004)
Complete middle school - incomplete college		0.006***	-0.006*	-0.010**		0.006***	-0.006*	-0.010**
		(0.001)	(0.004)	(0.004)		(0.001)	(0.004)	(0.004)
College degree or more		0.020***	0.006	0.004		0.020***	0.006	0.004
		(0.002)	(0.005)	(0.006)		(0.002)	(0.005)	(0.006)
Age group (18 to 24 omitted)								
25 to 29		-0.039***	-0.035***	-0.039***		-0.039***	-0.035***	-0.040***
		(0.010)	(0.012)	(0.012)		(0.010)	(0.012)	(0.012)
30 to 34		-0.056***	-0.043***	-0.047***		-0.056***	-0.043***	-0.047***
		(0.010)	(0.012)	(0.012)		(0.010)	(0.012)	(0.012)
35 or more		-0.070***	-0.044***	-0.047***		-0.070***	-0.044***	-0.047***
		(0.010)	(0.012)	(0.013)		(0.010)	(0.012)	(0.013)
Firm size dummies	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Region dummies	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Sector dummies	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Worker fixed effects	No	No	Yes	Yes	No	No	Yes	Yes
R^2	0.006	0.009	0.007	0.008	0.006	0.009	0.007	0.008
N	384,000	384,000	384,000	329,250	384,000	384,000	384,000	329,250

* All models include a constant term. Wage growth is calculated as $\ln w_t - \ln w_{t-1}$. Density is measured as the number of workers per km^2. Standard errors are presented in parentheses. Significance levels:* p<0.10, ** p<0.05, *** p<0.01.
Source: Author's own calculations.

The second set of results regarding the estimation of dynamic agglomeration advantages comprises regression models inspired in Equation 16, with the individual wage growth between t and $t-1$ as the dependent variable and a set of individual characteristics in t as explanatory variables. In this specification, it is possible to control for unobserved worker heterogeneity in relation to dynamic city-size earnings premium (Models 7 and 8 of Table 9).

The urban wage growth premium is defined as the relation between local density and individual wage growth. From Model 1 in Table 9, there is an indication of convergence of wages between cities according to their density. In fact, as aforementioned, wages are higher in denser cities. Therefore, a negative coefficient for higher density cities in relation to wage growth represents the fact that wages grow faster in less dense cities. The inclusion of additional observed controls at the individual level strengthens this effect (Model 2). However, the estimation with workers fixed effects (Model 3) reduces the significance of the coefficients related to density. Model 4 has the same specification as Model 3, but now the years in which workers have moved are excluded from the sample. Therefore, migrants stay in the sample, but their wage growth is calculated only between years in which they stay in the same place. The main results do not change significantly from Model 3 to Model 4.

Models 5 to 8 follow the same specifications of Models 1 to 4, but they also include previous work experience by city density levels. Focusing the analysis in Model 8, the dynamic wage growth premium obtained from working in other cities is positive and significant, and it is also higher for the experience obtained in denser cities. In this case, the static wage growth premium becomes non-significant. One additional year working in cities with density higher than 20 generates a variation of 0.7% in individual wage growth, while the experience accumulated in cities with density lower than 5 results in a variation of 0.4% in individual wage growth.

This analysis can be complemented by the evaluation of how this previous experience may have heterogeneous effects over wage growth according to the place the individual is currently working. Therefore, in Table 10, following the same preferred specification as before (Model 4 of Table 9, with workers fixed effects and excluding moving years), previous experience accumulated in cities with density lower than 5 have a positive and significant effect only in cities with similar density level or cities with high density level (20 or more). In addition, the experience accumulated in cities with medium density level is relevant to explain wage growth only in cities with the same or higher density levels. Finally, the experience accumulated in denser cities has a positive and significant effect over wage growth only in high-density cities. Once again, static agglomeration effects over wage growth become non-significant.

Table 10. Regression for the Urban Wage Growth Premium with Previous Experience and Current Workplace (dependent variable is the individual wage growth in one year), 2001 to 2008

	OLS Model 1	OLS Model 2	FE Model 3	FE Model 4
Density >= 5 & density < 20	-0.025***	-0.028***	0.007	-0.044
	(0.008)	(0.008)	(0.033)	(0.050)
Density >= 20	-0.029***	-0.030***	-0.011	-0.059
	(0.007)	(0.007)	(0.035)	(0.057)
Experience in cities with density < 5 (in years)*Now in city with density < 5	-0.029	-0.039	0.001	0.004***
	(0.038)	(0.038)	(0.001)	(0.001)
Exp. in cities with density < 5 (in years)*Now in city with 5 <= density < 20	-0.026	-0.036	0.000	0.008
	(0.038)	(0.038)	(0.003)	(0.005)
Exp. in cities with density < 5 (in years)*Now in city with density >= 20	-0.025	-0.034	0.008**	0.014**
	(0.038)	(0.038)	(0.004)	(0.006)
Exp. in cities with 5 <= density < 20 (in years)*Now in city with density < 5	-0.022	-0.032	0.013***	0.009
	(0.038)	(0.038)	(0.004)	(0.006)
Exp. in cities with 5 <= density < 20 (in years)*Now in city with 5 <= density < 20	-0.026	-0.036	0.004***	0.006***
	(0.038)	(0.038)	(0.001)	(0.001)
Exp. in cities with 5 <= density < 20 (in years)*Now in city with density >= 20	-0.025	-0.036	0.006***	0.009**
	(0.038)	(0.038)	(0.002)	(0.003)
Exp. in cities with density >= 20 (in years)*Now in city with density < 5	-0.027	-0.037	0.006	0.005
	(0.038)	(0.038)	(0.004)	(0.006)
Exp. in cities with density >= 20 (in years)*Now in city with 5 <= density < 20	-0.027	-0.037	0.002	0.003
	(0.038)	(0.038)	(0.002)	(0.004)
Exp. in cities with density >= 20 (in years)*Now in city with density >= 20	-0.026	-0.037	0.004***	0.007***
	(0.038)	(0.038)	(0.000)	(0.000)
Education attainment (incomplete primary school omitted)				
Comp. primary school - incompl. middle school		0.002	-0.006**	-0.010***
		(0.001)	(0.003)	(0.004)
Complete middle school - incomplete college		0.006***	-0.006*	-0.010**
		(0.001)	(0.004)	(0.004)
College degree or more		0.020***	0.005	0.004
		(0.002)	(0.005)	(0.006)
Age group	No	Yes	Yes	Yes
Firm size dummies	No	Yes	Yes	Yes
Region dummies	No	Yes	Yes	Yes
Sector dummies	No	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Worker fixed effects	No	No	Yes	Yes
R^2	0.0063	0.0089		
R^2 within			0.0073	0.0082
R^2 between			0.0055	0.0013
N	384,000	384,000	384,000	329,250

* All models include a constant term. Wage growth is calculated as $\ln w_t - \ln w_{t-1}$. Density is measured as the number of workers per km^2. Standard errors are presented in parentheses. Significance levels:* p<0.10, ** p<0.05, *** p<0.01.
Source: Author's own calculations.

These results indicate that previous experience has a relevant impact on wage growth only in cities with at least the same density level. Therefore, dynamic advantages are expected to be more relevant for migrants who move to denser cities. It is noteworthy that Model 4 of Table 10 excludes individual unobserved characteristics that may affect wage growth. Therefore, the process of sorting of high-skilled workers into denser cities is not affecting these results. Finally, there is no evidence of a pure wage growth effect once controls for previous experience are included in the specification (coefficients for the local density level are not significant in Model 4). This result is in accordance to D'Costa and Overman (2014).

In conclusion, individual wage level is expected to be more affected by previous work experience in less dense cities. However, future wage growth will be higher for individuals who migrate to higher density cities or for those who stay in cities of at least the same density group. This result indicates that apart from the sorting process, workers who move to bigger cities are expected to obtain higher salaries in the future.

7. Conclusion

This study has aimed to investigate static and dynamic advantages of working in agglomerated areas. The literature on agglomeration economies highlight the importance of the process of sorting of high-skilled workers into larger cities, which is usually captured by the inclusion of individual fixed effects. There are two main problems with this strategy (GROOT; DE GROOT; SMIT, 2014a): this sorting process becomes a black box (it is not actually explained or understood), and the variation to estimate agglomeration effects will come solely from migrants.
Therefore, following the most recent literature, a migration model was estimated for Brazil with data from RAIS-MIGRA.

The main conclusions from this analysis were that high-educated workers are more likely to migrate to work in another place, as well as individuals living in less dense cities. Furthermore, tenure at work and the time working in a certain REGIC area reduce the probability of outmigration. Therefore, the way incentives are distributed in space makes it more likely that regional inequalities will increase, as cities with a higher percentage of skilled workers will attract more of these high-qualified individuals. In addition, second migration seems to reinforce the effects of first-time migration. Workers are positively selected to migrate to a third place, but those who decide to return to the original REGIC area are less qualified, with lower salaries before the second migration in comparison to individuals who decide to stay).

Finally, the analysis of the correlations of wages in the origin in $t-1$ with the migration decision in t indicate that workers who have decided to migrate to a dense city in first-time migration (10 or more workers per squared kilometres) used to receive a higher wage in the origin. Even workers who have migrated to a less dense city (less than 10 workers per squared kilometres) received in average a higher wage than stayers did.

These results provide a sign that workers self-select into more agglomerated areas, and this process happens both in first as well as second-time migration decisions. In addition to that, the estimation of static agglomeration economies indicate that the inclusion of workers fixed effects capture most of the effect of local density over wages. Therefore, the sorting process explains a large share of the urban wage premium.

The inclusion of previous work experience in different cities aims to capture dynamic agglomeration advantages. In this setting, static agglomeration advantages are no longer significant. The main conclusion is that when years of previous experience are combined with the current place of work, individuals working in less dense cities who had previous experience in denser areas will benefit the most from these gains.

Moreover, the estimation of dynamic agglomeration advantages controlling for the possible worker heterogeneity in these gains indicates that previous experience has a relevant impact on wage growth only in cities with at least the same density level of the current place of work. Therefore, dynamic advantages are expected to be more relevant for migrants who move to denser cities.

In conclusion, workers who move to bigger cities are expected to obtain higher salaries in the future. This effect is supposedly net of the sorting process, as individuals fixed effects are controlled for. Combining these results with the conclusions from migration models, large cities attract high-skilled workers and provide them the conditions to obtain higher wage growth over time. Consequently, urban agglomeration economies are likely to increase wage inequality between cities. This process reinforces the attraction of more qualified workers to large urban areas, which will increase local productivity in these places.

Even though this movement towards spatial concentration of skilled workers could be expected to increase congestion costs in big cities, two other processes are supposed to happen at the same time. First, as noted in the main migration results, less skilled workers are more likely to return to their original REGIC areas, creating an outmigration flow. Furthermore, cities are reinventing themselves to bear with a higher demand for public services and land. Cities can become more productive, as it has already happened over time. Maybe new ways of organising work relations in space and provide public services with higher productivity could be possible solutions to face the increasing human flows to large cities.

Appendix

Table A.1 - Definition and Source of the Main Variables Considered in the Model

Variables	Definition	Level	Data source
Hourly wage	Monthly wage received in December divided by 4.3 times the number of weekly hours in the contract.	Individual	RAISMIGRA microdata
Age	Age at the end of the year.	Individual	RAISMIGRA microdata
Tenure	Number of years working in the same firm	Individual	RAISMIGRA microdata
Education attainment	Less than 8 years of schooling, 8 to 10 years of schooling, 11 to 14 years of schooling, 15 years of schooling or more	Individual	RAISMIGRA microdata
Firm size	Size of the firm in which the individual is working: up to 4 employees, 5 to 9, 10 to 19, 20 to 49, 50 to 99, 100 to 249, 250 to 499, 500 to 999, 1,000 employees or more.	Individual	RAISMIGRA microdata
Labour density in the formal sector	Total employment divided by the area (in km^2).	REGIC	RAIS - aggregated data
Area	Area in km^2.	REGIC	IPEADATA
Sector	manufacture; food and accommodation; transport and communication; finance, insurance, pensions and other services; real estate, rents and services to companies; education; health and social services; personal services and other sectors.	Individual	RAISMIGRA microdata
Macroregion	North, North-east, South-east, South, Centre-West	REGIC	
1st time migration	Dummy that equals one in the year the individual has moved for the first time since 1995	Individual	RAISMIGRA microdata
2nd time migration	Dummy that equals one in the year the individual has moved for the second time since 1995, provided he moved once.	Individual	RAISMIGRA microdata

Source: Authors' own calculations.

References

ANDERSSON, M.; KLAESSON, J.; LARSSON, J.P. The sources of the urban wage premium by worker skills: spatial sorting or agglomeration economies? **Papers in Regional Science**, Vol. 93, N. 4, p. 727-747, 2014.

BEHRENS, K.; DURANTON, G.; NICOUD, F.R. Productive cities: sorting, selection and agglomeration. **Journal of Political Economy**, Vol. 122, N. 3, p. 507-553, 2014.

BERRY, C.R.; GLAESER, E.L. The divergence of human capital levels across cities. **Papers in Regional Science**, Vol. 84, N. 3, p. 407-444, 2005.

CAMAGNI, R.; CAPELLO, R. CARAGLIU, A. Agglomeration economies and urban growth. **AISRe – ABC Workshop, The Frontier of the debate in Regional and Urban Economics**, available in http://www.aisre.it/images/call_for_paper/Workshop_dottorandi_ABC/03_Capello.pdf and accessed in 15/07/2015, 2015.

CARLSEN, F.; RATTSO, J.; STOKKKE, H. Education, experience and dynamic urban wage premium. **Norwegian University of Science and Technology Working Paper**, 2013.

CICCONE, A., HALL, R.E. Productivity and the density of economic activity. **The American Economic Review**, Vol. 86, N. 1, p. 54-70, 1996.

COMBES, P.P., DURANTON, G., GOBILLON, L. Spatial wage disparities: sorting matters! **Journal of Urban Economics**, Vol. 63, p. 723-742, 2008a.

COMBES, P.P., DURANTON, G., GOBILLON, L. The identification of agglomeration economies. **Journal of Economic Geography**. Vol. 11, p. 253-266, 2011.

COMBES, P.P., DURANTON, G., GOBILLON, L., ROUX, S. Chapter 1. Estimating agglomeration economies with history, geology, and worker effects. In: GLAESER, E.L. **Agglomeration Economics**. Chicago: The University of Chicago Press, p. 15-66, 2010.

COMBES, P.P., MAYER, T., THISSE, J.F. **Economic Geography**. Princeton: Princeton University Press, 1st Edition, 2008b.

COMBES, P.P.; GOBILLON, L. The empirics of agglomeration economies. In: DURANTON, G.; HENDERSON, J.V.; STRANGE, W.C. (Eds.) **Handbook of Regional and Urban Economics Vol 5A**, Amsterdam: North-Holland, 1st Edition, p. 247-348, 2015.

D'COSTA, S.; OVERMAN, H.G. The urban wage growth premium: sorting or learning? **Regional Science and Urban Economics**. Vol. 48, p. 168-179, 2014.

DE LA ROCA, J. Selection in initial and return migration: evidence from moves across Spanish cities. **NYU Working Paper**, N. 23, April 2015.

DE LA ROCA, J.; PUGA, D. Learning by working in cities. **Working Paper** Available in http://diegopuga.org/papers/esurban.pdf and accessed in 01/07/2015, 2014.

DURANTON, G. Agglomeration and jobs. In: FISCHER, M.M.; NIJKAMP, P. (eds.) **Handbook of Regional Science**. Spring-Verlag Berlin Heidelberg, p. 631-648, 2014a.

DURANTON, G. Growing through cities in developing countries. **The World Bank Research Observer Advance Access**, p. 1-35, Abril, 2014b.

DURANTON, G., PUGA, D. Micro-foundations of urban agglomeration economies. In: Henderson, J.V., Thisse, J.F. (Eds.) **Handbook of Regional and Urban Economics. Vol. 4.** Amsterdam: North-Holland, p. 2063-2117, 2004.

ELVERY, J.A. City size and skill intensity. **Regional Science and Urban Economics**. Vol. 40, p. 367-379, 2010.

FREGUGLIA, R.S.; GONÇALVES, E.; SILVA, E.R. Composition and determinants of the skilled out-migration in the Brazilian formal labour market: a panel data analysis from 1995 to 2006. **EconomiA**, Vol. 15, p. 100-117, 2014.

FREGUGLIA, R.S.; PROCÓPIO, T.S. Efeitos da mudança de emprego e da migração interestadual sobre os salários no Brasil informal: evidências a partir de dados em painel. **Pesquisa e Planejamento Econômico**, Vol. 43, N. 2, p. 255-278, 2013.

GLAESER, E.L. Learning in cities. **Journal of Urban Economics**, Vol. 46, p. 254-277, 1999.

GLAESER, E.L**. Triumph of the City: How our greatest invention makes us richer, smarter, greener, healthier, and happier**. London, UK: Macmillan, 2011.

GLAESER, E.L., MARÉ, D.C. Cities and skills. **Journal of Labour Economics**, Vol. 19 N. 2, p. 316-342, 2001.

GLAESER, E.L.; RESSEGER, M.G. The complementarity between cities and skills. **Journal of Regional Science**, Vol. 50, N. 1, p. 221-244, 2010.

GORDON, I.R. Ambition, human capital acquisition and the metropolitan escalator. **Regional Studies**, Vol. 49, N. 6, p. 1042-1055, 2015.

GREENWOOD, M.J. Internal migration in developed countries. In: ROSEMBERG, M.R.; STARK, O. (Eds.) **Handbook of Population and Family Economics, Vol. 1B**, Amsterdam: North Holland, p. 647-720, 1997.

GROOT, S.P.T., DE GROOT, H.L.F. Estimating the skill bias in agglomeration economies externalities and social returns to education: evidence from Dutch matched worker-firm micro-data. **Tinbergen Institute Discussion Paper** TI 2014-088/VII, 2014b.

GROOT, S.P.T., DE GROOT, H.L.F., SMIT, M.J. Regional wage differences in the Netherlands: micro evidence on agglomeration externalities. **Journal of Regional Science**, Vol. 54, Issue 3, p. 503-523, 2014a.

IBGE – Instituto Brasileiro de Geografia e Estatística. **Divisão Urbano Regional**, 2013.

LUCAS, R. E. On the mechanisms of economic development. **Journal of Monetary Economics**, Vol. 22, p. 3-42, 1988.

LUCAS, R. E. B..Internal migration in developing countries. In: ROSEMBERG, M.R.; STARK, O. (Eds.) **Handbook of Population and Family Economics, Vol. 1B**, Amsterdam: North Holland, p. 721-798, 1997.

MELO, P.C.; GRAHAM, D.J. Agglomeration economies and labour productivity: evidence from longitudinal worker data for GB's travel-to-work areas. **SERC Discussion Paper** 31, October 2009.

MION, G., NATICCHIONI, P. The spatial sorting and matching of skills and firms. **Canadian Journal of Economics**, Vol. 42, N. 1, p. 28-55, 2009.

MOULTON, B.R. An illustration of a pitfall in estimating the effects of aggregate variables on micro units. **The Review of Economics and Statistics**. Vol. 72, N. 2, p. 334-338, 1990.

PAPAGEORGIOU, T. Worker sorting and agglomeration economies. **McGill University Working Paper**, Sep. 2014.

PUGA, D. The magnitude and causes of agglomeration economies. **Journal of Regional Science**, Vol. 50, N. 1, p. 203-219, 2010.

RAMALHO, H.M.B.; QUEIROZ, V.S. Migração interestadual de retorno e autosseleção: evidências para o Brasil. **Pesquisa e Planejamento Econômico**, Vol. 41, N. 3, p. 369-396, 2011.

RATTSO, J.; STOKKE, H.E. Migration and dynamic agglomeration economies: regional income growth in Norway. **Norwegian University of Science and Technology Working Paper**, Feb. 2011.

ROSENTHAL, S.S., STRANGE, W.C. Chapter 49 – Evidence on the nature and sources of agglomeration economies. In HENDERSON, J.V., THISSE, J.F. (Eds.) **Handbook of Regional and Urban Economics. Volume 4**. New York: Elsevier, p. 2119-2171, 2004.

SANTOS-JUNIOR, E.R.; MENEZES-FILHO, N.A.; FERREIRA, P.C. Migração, seleção e diferenças regionais de renda no Brasil. **Pesquisa e Planejamento Econômico**, Vol. 35, N. 3, p. 299-331, 2005.

TAVEIRA, J. G.; ALMEIDA, E. Os determinantes regionais da atração do migrante qualificado. **Análise Econômica**, Ano 32, N. 69, p. 199-224, 2014.

WHEELER, C.H. Cities and the growth of wages among young workers: evidence from the NLSY. **Journal of Urban Economics**. Vol. 60, p. 162-184, 2006.

YANKOW, J.J. Why do cities pay more? An empirical examination of some competing theories of the urban wage premium. **Journal of Urban Economics**, Vol. 60, p. 139-161, 2006.