Agglomeration economies and college graduates job matching in Brazilian labor market

Adelson Santos Silva – UFRPE / UAST. E-mail: adelson.silva@ufrpe.br
Gervásio Ferreira dos Santos – PPGE-UFBA. E-mail: gervasios@ufba.br
Roberta de Moraes Rocha – PPGECON-UFPE-CAA. E-mail: Roberta_rocha_pe@yahoo.com.br
Diana Lúcia Gonzaga da Silva – PPGE-UFBA. E-mail: dlgsilva@ufba.br

Abstract

The aim of the paper was to analyze the relationship between agglomeration and college graduate job matching in Brazil. The research contributes by addressing the causal relationship between the labor market size and the matching of qualifications, considering different agglomerations in a large and heterogeneous geographical area. A panel data of formal college graduated workers in productive age in the period from 2006 to 2014 was used to investigate the influence of sorting and unobserved regional attributes on matching into more urbanized areas. Results showed that denser markets leverage better worker-job pairs. In addition, results on individual and agglomeration fixed effects showed that the influence of urban density on matching is inflated by the sorting of skills and regional attributes. Analogously, the urban wage premium was not driven by better matching in more urbanized environments. Similarly, the contribution of matching to productivity remains stable regardless of the size of the labor market.

Key-Words: Job Matching; Agglomeration; Urban Wage Premium

Resumo

O objetivo do artigo foi analisar a relação entre aglomeração e o *matching* de emrprego para trabalhadores qualificados no Brasil. A pesquisa contribui ao abordar a relação de causalidade entre o tamanho do mercado de trabalho e a correspondência de qualificações, considerando diferentes aglomerações em uma grande e heterogênea área geográfica. Um painel de dados de trabalhadores com nível superior, em idade produtiva, no período de 2006 a 2014, foi usado para investigar a influência do *sorting*, bem como de atributos regionais não observados sobre o *matching* de ocupação e qualificação em áreas mais urbanizadas. Os resultados sugerem que mercados mais densos alavancam melhores pares trabalhador-emprego. Além disso, os resultados dos modelos com efeitos fixos individuais e de área de mercado de trabalho mostraram que a influência da densidade urbana sobre o matching é inflada pela concentração de indivíduos mais habilidosos e de heterogeneidades regionais não-observadas. Por outro lado, os resultados para as equações de salários revelaram que o prêmio salarial urbano não é impulsionado por um melhor *matching* em ambientes mais urbanizados. Da mesma forma, a contribuição do *matching* para a produtividade permanece estável, independentemente do tamanho do mercado de trabalho.

Palavras-Chave: Matching; Aglomeração; Prêmio Salarial Urbano

JEL: I21; J24; R23

1. Introduction

The urban literature points that the largest cities are more productive (Glaeser; Maré, 2001; Combes *et al*., 2008; Korpi; Clark, 2019), because denser markets favor the sharing of

inputs and knowledge, the learning by externalities of human capital and the better matching of workers and firms (Duranton; Puga; 2004; Combes *et al*., 2010). Similarly, the intensity that these factors are spatially distributed also favors the attraction of firms and workers.

To Brazil, studies have shown the existence of a wage premium and suggest a positive relationship between the skills concentration and the productivity of workers in the largest cities (Servo and Azzoni, 2002; Rocha *et al.*, 2011). These studies are consistent with matching-based theories of urban agglomerations and point to a greater diversity of occupations and opportunities in the densest regions (Abel; Deitsz , 2014, Berlingieri, 2018). However, while the magnitude of agglomeration economies is well documented in Brazil, the identification of their microeconomic sources is still a challenge and, therefore, little progress has been made in understanding them.

In this work, we focus on matching as a microeconomic source of agglomeration economies. We start from the prerogative that denser environments facilitate matching between workers and firms, either through reduced job search costs or through greater diversity of career opportunities. Indeed, there is robust international evidence suggesting that greater urban density contributes to an increased probability of matching (Abel;Deitsz, 2014, Berlingieri, 2018). However, studies on the effects of city size on matching are still scarce. Most of them actually ignore how regional labor market conditions affect these outcomes (Robst,2007, Shimmer, 2007, Bleakely;Lin, 2012).

Recently, Abel and Deitz (2014) addressed these limitations proposing the *College Degree Match* as a matching measure for university graduates based on how well their jobs match their formal education. Following this approach, we focus on a direct measure of matching between the worker's education level and the job educational requirements. This measure is calculated based on the proportion of college graduates within each occupation to determine whether the college graduate is working in an employment that requires a college degree. Thus, we take use of a measure of vertical matching to identify the effect of urban density on the job matching likelihood.

This paper uses data from the Annual Social Information Report of the Brazilian Ministry of Economy (*RAIS/MTE*), an administrative database of worker and employer, which are compulsorily declared every year by companies in the formal labor market. The empirical analysis proposes to measure how much denser labor markets improve the chances of matching among college graduates. Linear probability models were estimated, based on the determinants of job matching for qualified workers in Brazilian agglomerations. The identification strategy proposed to deal with the sorting of unobserved heterogeneities that result from the attraction and concentration of the most productive workers in denser local labor markets. To control reverse causality, we estimate and compare models using two different instrumental variables. We apply a Bartik instrument to simulate the effect of demand shocks on urban density. On the other hand, considering that the literature has defended the use of historical instruments, as robustness check, we also use the population in 1910 as an instrumental variable.

We found that larger local labor market positively influences the vertical matching for college graduates, even when controlled by a set of observed characteristics of workers. However, when estimating the fixed effect models, the estimates systematically differ from baseline results. The results showed that agglomeration improves the chances of matching for the most skilled workers. Likewise, workers with the same educational background are exposed to greater competition for jobs in denser labor markets. The specifications with simultaneous fixed effects for individuals and regions suggest that the agglomeration effects on job matching are mainly related to better interactions of individual skills and the characteristics of local productive activity.

Finally, we investigated the contribution of job matching to graduates urban wage premium and found that matching increases individual-level wages. Thereby, evidences showed that job

matching modestly contributes to the wage premium. Thus, matching would relevant to explain differences in productivity among workers. However, our findings support the idea that the transmission of these effects fundamentally occur through other channels.

This research represents the first effort to isolate the causal effects of agglomeration on job matching in Brazil. In addition, it also provides a first attempt to control simultaneously unobserved individual and regional characteristics in estimating the effects of labor market size on job matching for qualified workers. By addressing fundamental identification issues and using a large panel database, the study adds new evidence to the literature on matching and allows to understand how heterogeneous labor markets influences the absorption of workers' qualifications in Brazil. Furthermore, this work represents an important step towards adding new evidence to the literature on agglomeration in developing countries, with a special focus on matching. It is noteworthy that the heterogeneities inherent to these countries print different characteristics in their labor markets and, therefore, influence the patterns of labor absorption and use of workers' productive characteristics. In these terms, different results are likely when compared to developed countries.

2 Evidences of Matching in Brazil

The urban economics literature consider that the largest labor markets concentrate most of the skilled workers (Glaeser; Maré, 2001, Abel and Deitz, 2011, Boualam, 2014). The trend to concentrate skills reflects lower transport costs or economies of scale in favor of biggest cities, which results in a greater concentration of productive firms (Marshall1890, Krugman, 1991, Venables, 2002). The greater diversity of employment opportunities in these cities may signal advantages to the concentration of workers in these locations (Abel and deitz, 2012, Boualam, 2014, Berlingieri, 2018). Thus, the agglomeration of qualified workers may result from a more complex sectoral structure of economic activity, which has dynamic effects on the occupational composition of the labor market. Therefore, these workers should be able to absorb different levels of human capital.

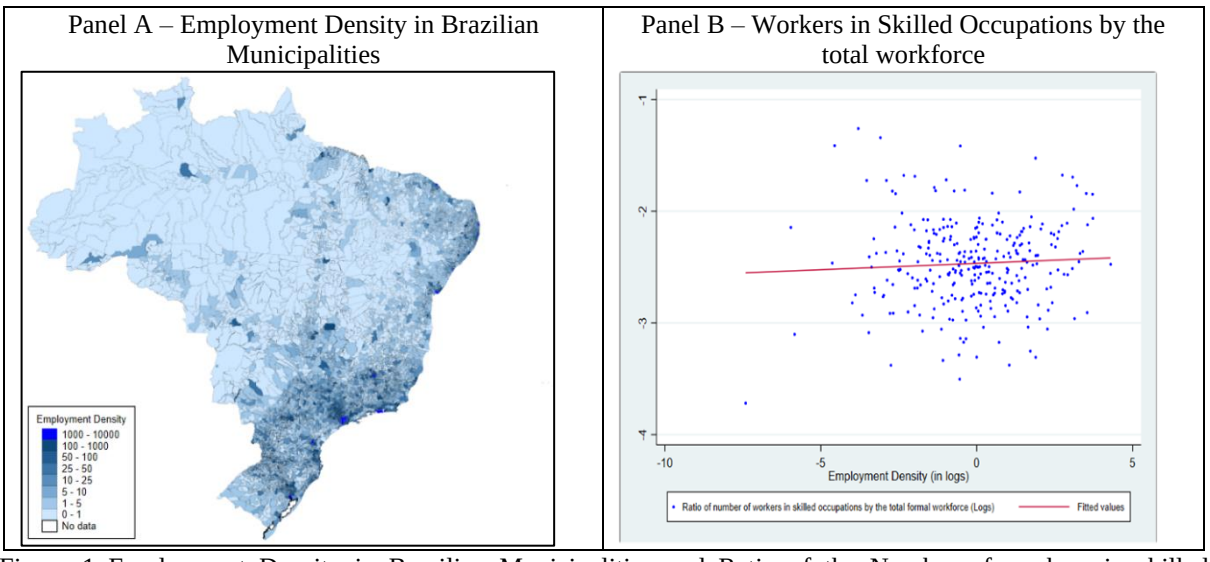

Figure 1 Employment Density in Brazilian Municipalities and Ratio of the Number of workers in skilled occupations by the total workforce.
Source: Authors based on Rais 2010-2014.

The Panel A in Figure 1 presents the employment density in Brazilian municipalities. The majority of cities are small labor markets and that the high-density urban centers are concentrated in the Center-South and in portions of the Brazilian coast. Recent evidence points to considerable heterogeneity in the regional distribution of employment in Brazil (Rocha *et*

al., 2011. Silva, 2017). This heterogeneity implies different job opportunities, especially for more qualified workers. Panel B reiterates this by demonstrating that the ratio of the number of workers employed in occupations with higher skill requirements grows with employment density. For this reason, the largest cities foster greater opportunities for absorbing skilled labor. As a consequence, regional inequalities may be the result of the heterogeneity in distribution of employment and in the configuration of the regional labor market. The Panel C in the Figure 2 shows that the total variation in the number of graduate job matches associated to employment density was 35.48%. The positive relationship between the log of the number of matches and the log of employment density reinforces the findings of Abel and Deitz (2014) and Berlingieri (2018), that find better matching in the denser markets.

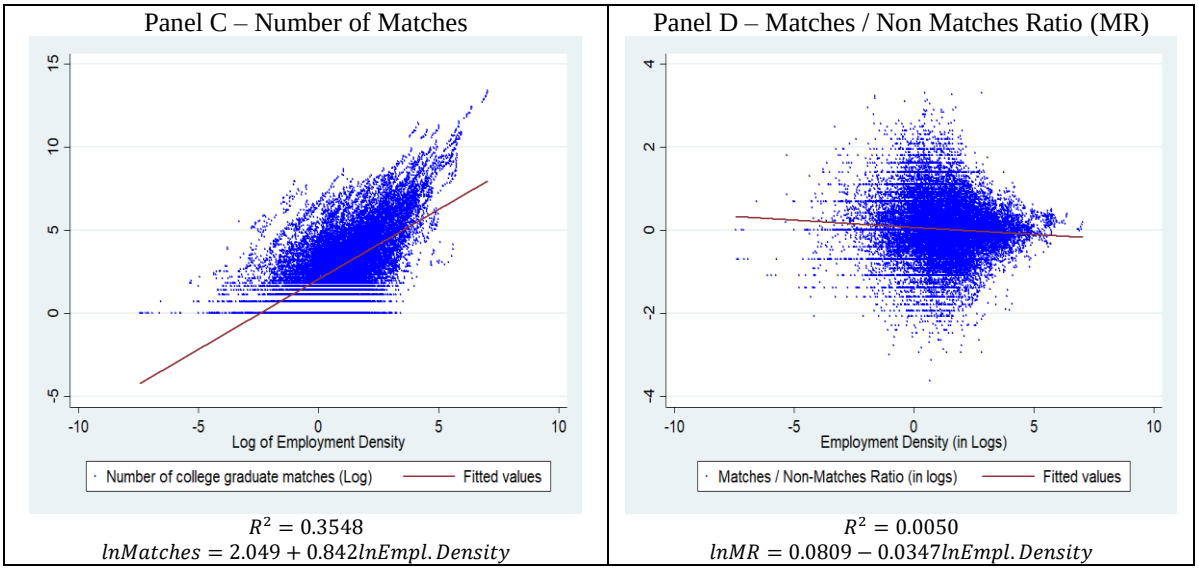

Figure 2 – Distribution of matches and urban density in Brazil
Source: Authors based on RAIS 2010-2014.

On the other hand, as can be seen in the panel (D), the elasticity of the match and non-match ratio was about -3.5%. Such evidence is comparable to the results of Reis (2012), which pointed to an increase in over-education in the largest Brazilian regions. Despite the size of the labor market, it's likely that such markets are not dynamic enough to generate more complex jobs that are able to absorb skilled workers, fostering bad matches.

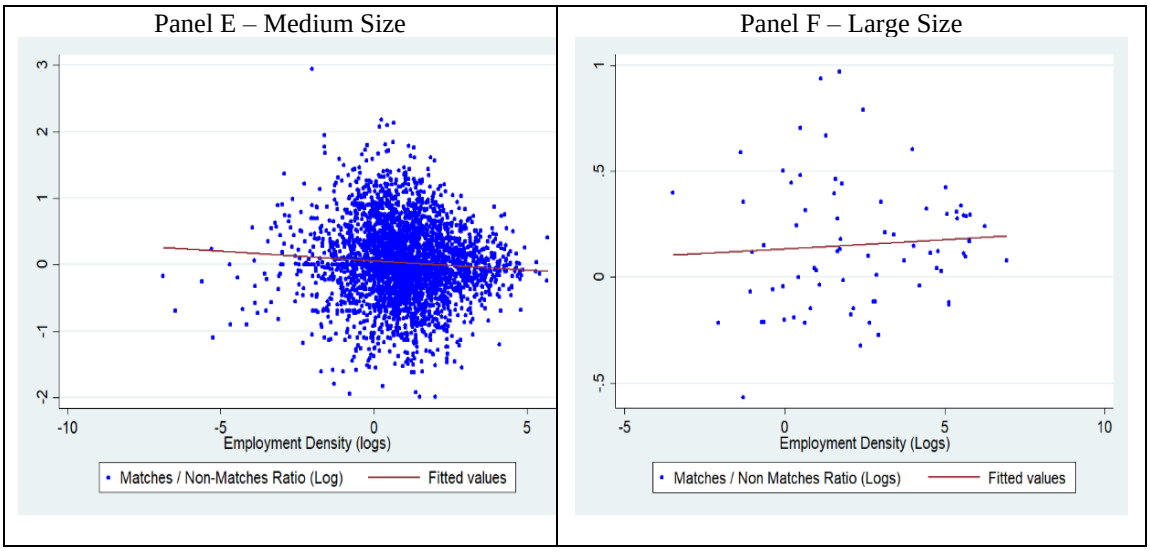

Figure 3: Employment density and matches/non matches ratio in Brazilian Agglomerations

Source: Authors based on Rais 2010-2014.

Figure 3 shows that the employment density improves the matches/non-matches ratio in large population agglomerations, but with an opposite association on the medium ones. Thus, the size of the labor market positively influences on the job matching only in cases where they have an appropriate industrial structure for the attraction of skills. The largest agglomerations seem to have a better organization of economic activity, due to greater breadth of the consumer market and the existence of pronounced agglomerative effects. In addition, the large agglomerations also concentrate greater sectorial and occupational specialization. This specialization enhances the absorption of skills and determines the better allocation of workers with higher education. In summary, the effects of density on the matches/non-matches ratio depend on the capacity of the local economy to generate economic activities that demand and absorb the stock of individual human capital.

3. Agglomeration and job matching: Empirical Strategy

3.1 Data and Variables

This paper uses the Annual Social Information Report of the Brazilian Ministry of Economy (*RAIS/ME*) to generate an administrative database of worker and employer, which are compulsorily declared every year by companies in the formal labor market. These data allow the geographic, sectorial and occupational monitoring of formal workers over time. Data were restricted to college graduated workers in productive age, from 18 to 65 years in the period from 2006 to 2014. We excluded the public sector workers due to the different dynamics in the hiring process in comparison to private sector. Finally, workers with inconsistent information of location, gender and born date also were excluded. Thus, we have an unbalanced panel with 28,888,353 observations relating to circa 4,588,628 workers.

The individual control variables include age, squared age, gender and job tenure, measured in months in the same employment. We also include dummies of economic activity sector, according to the classification of economic activities of the Brazilian Institute of Geography and Statistics (IBGE). As well, four dummies of firm size were added, following the classification of Inter-Union Department of Statistics and Socioeconomics Studies (DIEESE). Small firms are those with 10 to 49 workers in Service and Commerce Sectors and 20 to 99 in Manufacturing; mediums are those with 50 to 99 workers in Service and Commerce, and 100 to 499 in Manufacturing and big firms are those with above 100 workers in Service and Commerce, and above 500 in Manufacturing.

We use the Brazilian Classification of Occupations (CBO-2002), a document of the Brazilian Ministry of the Economy that catalogues about 2529 occupations and codify their titles and contents, based on the methodology of Developing a Curriculum (DACUM), which uses supervised discussions to guide the systematic organization of information related to each occupation. The CBO contains the level of expected education for each occupation (Maciente, 2012), which might be used to organize information regarding the level of skills and required competencies from workers. Moreover, CBO corresponds to the Uniform International Classification of Occupations (CIUO-88) of the International Labor Organization (ILO). The CIUO-88 system defines 4 increasing levels of skills required by occupations. Those with level 1 demand less qualification, while level 4 implies in most demanding qualified labor. The CIUO-88 model for the Brazilian labor market considers an interpretation of "skills" whose level is punctuated more strongly by the complexity of the activities performed by workers than by the level of education (CBO-MTE, 2017).

The literature proposes the matching variable from the self-assessment of worker about the educational requirements for the job (Abel; Deitz, 2012, Berlingieri, 2018). However, the absence of surveys in Brazil makes this type of treatment unfeasible. The limitation of identified training data also restricts the analysis of job matching. Brazilian labor databases do not monitor careers, while education databases do not provide identification of individuals. To overcome these issues, we adopt a vertical matching measure based on Boualam (2014) and Abel; Deitz (2014). It is estimated by the comparison of the highest level of worker's education and the highest education level required by the occupation in which he is employed. Because of the focus on qualified individuals, the strategy consisted of identifying the occupations with the highest capacity to attract workers with higher education. So, we built a variable based on the proportion of college graduate within each occupation.

Following Abel and Deitz (2014) and Berlingieri (2018), qualified occupations were defined as those with a proportion greater than or equal to 50%[1]. Database checks suggest that more than 90% of CIUO level 4 occupations, i.e., those with higher requirements for skilled workers, have more than 50% of graduates in the total employment. This group absorbs Science and Arts Professionals, whose higher education is a fundamental requirement for most workers. A dummy variable was built based on the match (Match=1, if a graduate is employed in a job whose proportion of workers with a college graduation is greater than or equal to 50%). Thus, matching was defined as the vertical correspondence of the worker's occupation and education. Identifying the regional capacity for absorbing formal human capital is an important advantage of the methodological approach proposed in this research. A bad matching, for example, is associated to the regional limitations of economic activity to generate complex occupations to absorb the qualification of the workforce.

The agglomerations employment density was calculated by the number of individuals employed per square kilometer and used as a measure of the labor market size to isolate the influence of urbanization economies. An agglomeration is defined as a concentration of municipalities with strong integration of population movements[2]. The IBGE identifies 294 urban agglomerations (arrangements) in the Demographic Census of the 2010. These agglomerations are composed of 938 municipalities, which concentrate about 56% of the Brazilian population. Based on the population ranges, the agglomerations were grouped: 189 small (population ≤ 100 thousand inhabitants); 81 mediums (100 thousand <population ≤750 thousand) and 24 large (population> 750 thousand inhabitants). Furthermore, single municipalities with a population of more than 100,000 inhabitants were also considered as agglomerations. Thus, we have 26 large agglomerations (2 isolated municipalities + 24 agglomerations) and 158 medium agglomerations (77 municipalities + 81 agglomerations).

The Brazilian municipalities underwent significant changes in administrative boundaries from 2006 to 2014. For this reason, we adopt the Minimum Comparable Areas (AMC-2013)[3] to define the aggregations of municipalities that were dismembered or created (IBGE). We have 4,256 AMC corresponding to 5570 municipalities. We made them compatible with agglomerations in order to maintain the temporal consistency of the analysis. Thus, 365 agglomerations were defined: 177 small (495 municipalities); 161 medium (596 municipalities); and 27 larges (301 municipalities). The potential effects of urban composition

[1] We tested 40% and 60%, but there were no systematic differences.

[2] The evidence of contiguity of urban areas, conurbation and intense commuting movements to work or study are factors that used to measure this integration.

[3] A minimum area corresponds to the junction or aggregation of areas from all municipalities of origin, considering the relevant inter-census period. In the case of municipalities that originate from more than one municipality, this implies aggregating the areas of all municipalities of origin. Comparable Minimum Areas (AMC) do not refer, therefore, to a political or administrative division, but simply to the aggregated area of the smallest number of municipalities necessary for the intertemporal comparisons to be geographically consistent (Ehrl, 2017).

concentration as well as differences in the economic performance of the regions also were controlled. Sectoral specialization of regional employment was used by the measure of Herfindhal-Hirschman Sectorial index, and the regional participation of high skilled occupations (CBO / RAIS). To calculate the hourly wage, the December average wage deflated by the Broad Consumer Price Index (IPCA) was used, considering 2014 as the base year. The index was calculated for the twelve most important metropolitan regions and extrapolated to the other corresponding states, according to proximity. The actual monthly wage was then divided by the number of monthly hours in order to obtain an hourly wage measure.

3.2. Econometric Approach

Based on the positive relationship between matching and agglomeration, the linear probability models were estimated to test the robustness of this relationship using a wide set of control variables, according to the following specification:

$$Pr(\mathcal{M}_{ijt} = 1) = \alpha + \beta Emp_{jt} + \delta Hop_{jt} + \theta Hst_{jt} + \gamma X_{ijt} + \delta F_{it} + \Gamma Z_j + \varepsilon_{ijt} \qquad (1)$$

In the equation (1), $\mathcal{M}$ is a dummy that takes value 1 if a graduated worker *i* works in a job that requires for a graduation in year *t*. The match is between college degree (actual level of qualification) and the job requirements, so that the measure refers to the vertical matching. *Emp* is the employment density of the working region *j* in the year *t* and Z_j are dummies of macroregion. The effects of urban specialization were controlled by employing Herfhindal-Hirschman for occupations ($\mathrm{Hop_{jt}}$) and industries (Hst_{jt}). To address the differences in worker characteristics, we include a range of individual level characteristics (X_{ijt}) that are expected to influence the likelihood of matching, such as age, sex and tenure. In addition, dummies for firm size and for industry (F_{it}), were included, since the characteristics of the firm should affect the matching results from an individual perspective. Large firms, for example, reflect a more robust production structure, based, above all, on greater economies of scale and better absorption of qualified labor.

Studies in Urban Economics show that biggest cities attract more skilled workers (Glaeser; Maré, 2001). There is an inherent self-selection in denser urban environments, which may be behind both the better quality of matches and increases in productivity. Thus, to control the bias of more skilled workers migration, we restrict the sample to individuals who remain in the same region, since empirical evidence points to characteristics more favorable to individuals who stay in more agglomerated labor markets (Glaeser; Maré, 2001; Andini, 2013; Boualam, 2014; Berlingieri, 2018). By imposing a restriction to non-movers, it is possible to control the influence of workers with a high skill load on the chances of matching. Moreover, it is possible to control the effects of regional skills transfer on the process of adjusting the supply and demand for skills in different labor markets.

Even controlling for non-movers, a residual source of endogeneity remains if skilled workers are likely to stay in cities (Glaeser; Maré, 2001, Gould, 2007, D'Costa; Overman; Puga, 2013). Skilled workers settle in large markets because they have a broader set of information, so that they know better the vacancies corresponding to their qualification. In addition, consumer amenities should make large cities more attractive to skilled people, even with positive cost of living differentials. Furthermore, if labor markets remain heterogeneous with each other, it is likely that unobservable attributes of the region will affect the sectorial and occupational configuration of the labor market as well as the worker's choice of location. As a consequence, the effect of employment density on graduates' job matching may be incorporating sorting and region effects, in order to generate biased estimates.

To address these sources of endogeneity we estimated the models by gradually including individual and regional fixed effects.

$$Pr(M_{ijt} = 1) = \alpha + \beta DensEmp_{jt} + \delta Hop_{jt} + \theta Hst_{jt} + \gamma X_{ijt} + \emptyset_i + \sigma_j + \varepsilon_{ijt} \quad (2)$$

Where $\emptyset_i$ and σ_j are individuals and regions fixed effects, respectively. By controlling for such effects, it is possible to identify the influence of the regional concentration of skills on matching. If more skilled individuals are attracted to the larger labor markets, part of the effects of density should reflect this greater concentration. Economic activity may be shaped by unobserved factors specific to location, such as proximity to infrastructure for the acquisition and disposal of products, institutional quality, availability of natural resources. The geographical disposition of these characteristics when directing productive opportunities and attracting firms must have some relationship with matching, especially if the local productive activity is strongly based on the use of these resources. At the same time, if local amenities, despite the higher costs of living in cities, attract more skilled workers (Venables, 2002), regional attributes should affect the likelihood and quality of correspondence in urban environments. The omission of these effects would imply skewed estimates of the effect of density on matching. Since a fundamental part of these characteristics is fixed over time, the inclusion of a fixed region effect should assist in the treatment of endogeneity associated with the omission of these variables.

3.3 Reverse Causality

Urban studies have warned of the potential reverse causality between urban density and local productivity. Higher wages and locations with improved matching tend to attract workers, which has effects on the supply of regional labor and, therefore, should affect the density of employment (Combes; Gobillon, 2015). Similarly, both firms and worker localization decisions tend, in part, to absorb the effect of local productivity shocks, and in that sense, employment density may be capturing such shocks. On the other hand, the relationship between job matching and agglomeration is expected to be endogenous due to the possibility of better job matching increasing the urban density by attracting workers and firms. To address the potential reverse causality, we re-estimate our models using instrumental variable.

To address the potential reverse causality, we use the sectoral employment Bartik as instrumental variable (Bartik, 1991). This instrument estimates the expected sectorial growth based on the growth of the respective sector, taking as reference the initial share of local employment in the sector.

$$Bartik_{agg,t} = \sum_s E_{s,agg,t_0} \left(\frac{E_{(n-agg),s,t} - E_{(n-agg),s,t_0}}{E_{(n-agg),s,t_0}} \right) \quad (3)$$

Where E_{s,arr,t_0} is the employment level, in sector s, in the arr arrangement and in the initial time t_0 (1995). The terms $\left(\frac{E_{(n-agg),s,t} - E_{(n-agg),s,t_0}}{E_{(n-agg),s,t_0}} \right)$represent national employment growth (n) of the sector (s) in each year (2006-2014), starting from the reference year (t_0), excluding the arrangement employment $(n - arr)$. We excluded individually the employment in each arrangement in the calculation of national growth. This instrument provides an estimate of the growth of local employment based on the growth of national sectoral employment. It predicts the growth of local employment originating from shocks of national demand, in view of the exogeneity of local shocks. Thus, employment growth induced by demand shocks would represent an indicator of the size of the national shock on the local economy. The Bartik version in this study was calculated from the growth of national employment in each of the 25

subsectors of Geography and Statistics Brazilian Institute-IBGE, available in the RAIS-MTE database (2006-2014).

The difference between the additional growth of local employment in relation to the predicted growth, based on the sectoral structure, would represent the share of employment explained by local productivity shocks. The instrument's exogeneity is guaranteed by excluding local employment, as it is possible that there is a high spatial concentration of a sector in one or a few regions, which would imply a strong correlation between local shocks and national employment (Silva, 2017). However, if the endogeneity is caused by contemporary local shocks, the historic distribution of population should not be correlated with current productivity. Urban density, in its turn, is determined by past patterns of population concentration. Thus, historical variables can be used as instruments (Combes *et. al*, 2008). To deal with this and to compare with bartik estimates, we re-estimated the model by applying log of population in agglomeration in year of 1910 as an instrumental variable for employment density.

Considering that the Brazilian agglomerations highlights the particularities of the process of economic formation in Brazil, it is likely that Bartik underestimates the endogenous relationship between matching and employment density. On the other hand, if the greater concentration of skilled workers in certain locations reflects more recent demand shocks, which reverberate in greater diversity and development of economic activity, the effect of endogeneity on matching may be overestimated in the model with the log of population in 1910. From the 2000s onwards, Brazil underwent significant urban transformation, including new trends in population attraction beyond the Rio-São Paulo axis. Thus, urban concentration does not absolutely reflect more remote population patterns, but rather the impulses of more recent demand policies.

4. Descriptive Statistics and Results

Table 1 presents the descriptive statistics of the main variables used in the analysis. The statistics point out to a greater number of graduate jobs matches into densest labor market. In the largest agglomerations about 54% of the graduates were employed in qualified occupations. Medium agglomerations also showed a slightly higher proportion of matches than non-matches. However, for the small agglomerations the proportion of matches was less than 50%. The log of hourly wage was about 31% higher in the densest agglomerations compared to the medium, and 64% higher compared to the smallest. Considering the differences in cost of living, the wage differential points to differences in productivity in favor of more agglomerated areas. Finally, the log of employment density is around 60% greater in the largest agglomerations compared to the medium-sized ones, and twice as large compared to small ones. Since the productive characteristics of both workers and firms are more concentrated in densest places, urban density explains both matching and wages (Combes *et al*., 2010, Abel; Deitz, 2012).

Table 1 also show the predominance of female workers, regardless of the agglomeration size. The average age of workers is slightly higher in large agglomerations. However, the higher standard deviation indicates greater variability in the age group of these workers. In relation to economic activity sectors, the highest incidence of matches was verified in the educational activities, about 77%. The Health and Social Service and Financial Intermediation sectors also presented high proportion of matches, 73% and 62%, respectively. Thus, the sectoral distribution of higher education employment matches might be related to the ability to generate complex occupations and the potential to absorb the worker's skills. The Transformation Industry and Trade concentrates the greatest number of workers and present the lowest

proportion of matches. These industries demand skills to handle engines, so that a higher education is not a fundamental requirement for the development of good practices in the execution of activities (MACIENTE, 2012). At the same time, the high concentration of graduate workers in these sectors reflects imbalances in the structure of absorption of qualified workers in the Brazilian labor markets.

Table 1. Summary Statistics

	Large Arrangements		Medium Arrangements		Small Arrangements		Others	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Dependent Variables and other main variables								
Matches	0,54	0,50	0,51	0,50	0,48	0,50	0,49	0,50
Real Wage	6141,09	6,872,50	4422,08	5173,83	3535,95	4113,57	3574,63	4283,23
Hourly Wage	171,79	324,85	130,76	259,10	104,60	199,26	100,23	156,21
Hourly Wage (log)	4,68	0,93	4,40	0,91	4,19	0,89	4,16	0,90
Employment Density (log)	5,96	1,01	3,75	1,15	2,61	1,25	1,92	1,37
Employment Density	571,89	410,77	68,15	63,52	23,86	22,42	13,93	16,58
Employment	3579249	2936798	104675,9	62492,68	16918,89	8121,46	10418,09	9634,04
Individual Characteristics								
Men	0,488	0,50	0,471	0,499	0,457	0,498	0,479	0,50
Age	36,47	9,76	35,82	9,62	35,72	9,71	35,34	9,58
Tenure	60,43	81,97	60,31	77,80	60,18	76,47	57,59	5,09
Industry								
	Obs	Percent	Obs	Percent	Obs	Percent	Obs	Percent
Agriculture	46673	0,23	59014	1,06	20617	2,84	135663	5,10
Mining	180758	0,91	137124	2,45	6214	0,86	58549	2,20
Transformation	2730817	13,73	1128329	20,19	154201	21,24	531326	19,98
Water and electricity	257823	1,30	46606	0,83	4568	0,63	22356	0,84
Construction	599907	3,02	131053	2,35	11535	1,59	49248	1,85
Trade	2344587	11,79	742184	13,28	118754	16,36	461735	17,36
Food and Accomodation	264747	1,33	76470	1,37	14818	2,04	27274	1,03
Transport and storage	1396503	7,02	239349	4,28	17449	2,40	76961	2,89
Financial	2633211	13,24	572694	10,25	86960	11,98	355893	13,38
Real Estate	3823056	19,22	536463	9,60	48018	6,61	164975	6,20
Security	117842	0,59	13917	0,25	4620	0,64	4160	0,16
Education	2337754	11,75	1000392	17,90	129507	17,84	416827	15,68
Health and Social Service	1434006	7,21	464166	8,31	56408	7,77	171616	6,45
Colective Services	1710149	8,60	439059	7,86	52280	7,20	182330	6,86
Home Services	439	0,00	187	0,00	37	0,01	197	0,01
International Organizations	9593	0,05	1222	0,02	36	0,00	56	0,00
Observations	19804295	100	5708967	100	714636	100	2659166	100

Source: Authors based on RAIS 2010-2014.

4.1 Agglomeration and Matching in Brazilian Labor Market: Baseline Results

The estimates of the Linear Probability Models are in Table 2. Column (1) presents gross estimates of the impact of urban density on college graduates job matching. The results show that an 1% increase in employment density increases 0.27% the chance of matching. The positive and significant coefficient at 1% indicate that denser environments promote a more adjusted interaction between worker and job. The better matching of skill requirements of the

vacancies with the individual characteristics is consistent with the theories of urban agglomeration based on matching (Helsey; Strange, 1990, Berliant, 2006, Duranton; Puga, 2008). Furthermore, the results are close to the findings of Rocha et al. (2011), Silva (2017) and Amarante *et al*. (2017), that denser Brazilian markets promote better use of individual productivity.

Table 2 – Impact of employment density on College Graduates' Job matching in Brazil

	(Panel LPM)	(Panel LPM)	(Panel LPM)	(Panel LPM)	(Panel LPM)
	(1)	(2)	(3)	(4)	(5)
Empl.Density	0.00273***	0.00228***	0.00285***	0.00323***	0.00247***
	(0.000097)	(0.000096)	(0.000095)	(0.000111)	(0.000115)
HHI Ocup					0.349***
					(0.00904)
HHI Setor					0.105***
					(0.00336)
Individuals	No	Yes	Yes	Yes	Yes
Firms	No	No	Yes	Yes	Yes
Geographic	No	No	No	Yes	Yes
HHI Occup.	No	No	No	No	Yes
HHI Industry	No	No	No	No	Yes
Year	Yes	Yes	Yes	Yes	Yes
_cons	0.430***	0.170***	0.118***	0.115***	0.0867***
	(0.000559)	(0.00216)	(0.00268)	(0.00273)	(0.00294)
Obs	28888353	28888353	28861282	28576647	28576647
R2	0.0002	0.0043	0.0944	0.0943	0.0938
Prob>F	0.0000	0.0000	0.0000	0.0000	0.0000

Source: Authors based on RAIS 2010-2014. Note: Standard errors are clustered at the individual level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The remaining four columns report the results from gradually including individual, firm and region characteristics. Since we have a homogenous sample in terms of education level, we include age, gender and tenure as the main individual control variables. Dummies variables for industry and for the firm size (main firm controls) were included. The analysis also complains dummies to the 5 Brazilian macro-region and Herfindahl-Hirschman index for Occupation and Industry as urban specialization proxy (Column 5). The estimates for the influence of individual characteristics on job matching showed that the chances of matching increase as the firm size grows. The findings also signal to a better job matching in most sectors in comparison to agriculture. Education and health and social services industries are the most prone to the graduate job match, with respective chances of 27% and 22%. However, for jobs in the Manufacturing sector the chances of matching to qualified workers are worse in Brazil. The apparent low complexity and the reduced dynamism of Brazilian industry might explain this result. This evidence is in accordance that presented by Rolnik and Klink (2011), which show a relative concentration of the more complex industrial segments in few cities in the Brazilian Southeast and South. Furthermore, results are in line with the idea that a great part of the Brazilian industry is specialized in the production of non-durable goods and, therefore, had a reduced demand for qualified labor.

The urban density can capture the effects of the sectorial and/or occupational composition of cities and the effects related to specialization. For this reason, two additional measures were implemented as regional controls (Column 5). The first is the Herfindahl-Hirschman index for occupational concentration. The use of this measure aimed both to identify how the specialization of regional employment influences on matches; and to point to the likely effects of the competition of workers on the probability of match. The second was the regional index of sectorial concentration, in order to gather evidence on the relationship between specialization of economic activity and matching.

The estimated coefficient for the Hirschman-Herfindhal index for occupations was 34% and significant at 1%, which demonstrates a greater influence of the specialization of regional employment on the probability of matching. Industrial Specialization increases the graduates

job match by up to 10%. The result is in line with Boualam (2014), which points to independence between agglomeration and specialization to explain first job matching for graduates in the French labor market. In this context, for the Brazilian case, the match between occupation and qualification seems to be more related to both labor market and economic activity structure than urban density. Similarly, the results suggest that workers can benefit from being located in specialized markets rather than by staying in more urbanized areas.

Based on the hypothesis of localization economies, specialized firms must locate in regions with greater endowment of qualified workers and easier access to inputs and raw materials (Marshall, 1890). In addition, empirical evidence points to a concentration of technology-intensive industries in large Brazilian urban centers (Silveira Neto, 2005, Maciente, 2012). Thus, the influence of sectoral concentration on the job match among graduates may reflect a labor pooling. Firms benefit from the greater concentration of workers. As a consequence, the workers have access to a diverse and dynamic range of qualified jobs.

Finally, after controlling for the influence of individual, firms and region characteristics, we found a coefficient of 0.00247, similar to that found in Abel and Deitz (2014). The results are in line with the theories of urban agglomeration based on matching. The denser labor markets foster greater productivity, since it improves the chances as well as the quality of matches (Helsey; Strange, 1990, Sato, 2001, Berliant *et al*., 2006, Abel; Deitz, 2012, Berlingieri, 2018). The evidence suggests that part of the regional differences in the absorption of qualification and in the labor market matching might be explained by agglomeration.

4.2 Individual and Regional Heterogeneities and the Determination of Matches

The models (1) and (2) of Table 3 reports the results for the pooled linear probability model (PLPM). As shown in Column (2) non-mover graduated workers are twice as likely to match to a qualified occupation. Based on the hypothesis that more skilled workers remain in the largest labor markets, where there are agglomeration economies, this result points to a skill sorting in favor of larger agglomerations.

Table 3 – Impact of employment density on college graduates' job matching: Addressing skills spatial *sorting*

	(PLPM)	(PLPM)	(Individual FE)	(Individual FE)	(Individual and Agglom.FE)
	All Workers	Non Movers	All Workers	Non Movers	All Workers
	(1)	(2)	(3)	(4)	(5)
Emp. Densi (Log)	0.00506***	0.0106***	-0.00335***	-0.00753***	-0.0193***
	(0.000120)	(0.000138)	(0.000215)	(0.00219)	(0.00291)
					0.201***
					(0.0215)
					0.103***
					(0.0106)
Individuals	Yes	Yes	Yes	Yes	Yes
Firms	Yes	Yes	Yes	Yes	Yes
Industries	Yes	Yes	Yes	Yes	Yes
Geographic	Yes	Yes	Yes	Yes	Yes
Years	Yes	Yes	Yes	Yes	Yes
_cons	0.281***	0.271***	0.148***	0.158***	0.151***
	(0.00307)	(0.00346)	(0.00807)	(0.0197)	(0.0197)
Obs	28888353	23319564	28576647	23319564	23919733
Prob>F	0.000	0.000	0.000	0.000	0.000
R^2	0.100	0.108	0.017	0.016	0.783

Source: Authors based on RAIS 2010-2014. Note: Standard errors are clustered at the individual level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The column (3) brings the estimates by individuals fixed effect. The coefficient for matching (-0,003) turns out to be negative and systematically differs from those found in the baseline results. It is statistically significant at 1%. In the column (4) we regress the model only for individuals who did not move to another agglomeration. At 1% the result (-0,007) shows that graduates' job matching in Brazil is strongly correlated with the unobservable skills. Such evidence also corroborates that a better matching in largest Brazilian urban centers derive from an improved information set resulting from the greater concentration of unobservable skills in these environments (Glaeser; Maré, 2001).

Competition for vacancies in the densest labor markets should lead to less chances of matching for the unskilled workers. Thus, the higher density can cause allocative mismatches in the structure of formal human capital absorption if the individual does not carry with him the appropriate set of skills that allows the most efficient and productive use of his qualification. In addition, if the region does not have a relevant concentration of individual observable and unobservable skills, the production structure will probably mirror this shortage, generating jobs that are also less complex and therefore less skilled.

Urbanization economies, in this context, seem to have little influence on graduates' job matching. Actually, the effect of the employment density on matching not only disappeared, but there was a significant change in the relationship between the variables. As found in Silva (2017), the productivity in large urban centers is shown to be more strongly correlated with the disposition of skills than by the density effect itself. Thus, the effect of labor market size on the job matches reflects, in fact, the interrelationship between the sorting and the qualification job matching. This relationship, in parallel, signals the market's ability to absorb the productive characteristics of workers (Duranton; Puga, 2008). Dense markets, in this sense, would not absorb skilled workers due to their size, but because the concentration of skills attracts more productive firms, resulting in the formation of more efficient pairs.

In line with the Marshallian hypothesis of location economies, matching as a source of agglomeration seems to reflect the way in which economic activity is organized. Likewise, if denser regions enhance better matching by concentrating more skilled workers, productivity gains are generated if the market specializes in activities whose sharing of labor and inputs is facilitated. Again, the evidence suggests that specialization is more relevant than urbanization to explain matching in the Brazilian labor market.

However, Berry and Glaeser (2005) suggested that the most skilled regions tend to internalize more skills over time, mainly due to the attraction of more qualified labor. Evidence even suggests that more talented individuals are incorporated more efficiently into markets specialized in more skilled and complex sectors of the economy (Krugman, 1991, Abowd et al., 1999). The organization of economic activity tends to reflect the stock of regional skills. On the other hand, the firms are located in places with a greater disposition of human capital and with additional advantages to the installation of productive plants. The result of this dynamic is the hiring of more talented workers. In this sense, the effect of the density on job matching can also reflect a regional component, including among other factors, the internal mechanisms for selecting more productive workers. Table 4 provides estimates of the effect of employment density on matching, considering multiple fixed effects.

The column (5) shows the estimates by considering simultaneous fixed effects for individuals and arrangements. The sign of the coefficient of employment density remained negative, but the drop in the chance of matching is six times greater if discounted individual and regional characteristics simultaneously compared to the individual fixed effect specification. In its turn, the occupational and sectoral specialization coefficients proved to be relevant, despite the loss of efficiency in the model for non-movers (5% only).

Estimates by fixed effects (Table 3) differ systematically from baseline regressions and reveal that graduates' job matching is strongly related with unobserved regional productive

attributes and worker's skills. In this context, workers with the same educational background in denser markets, with a greater supply of skills, are exposed to greater competition for vacancies. On the other hand, more skilled markets develop more skilled jobs and therefore attract workers with a better stock of observable and unobservable skills, consistent with findings of Rauch (1993).

Finally, these findings reveal that the largest agglomerations only enhance greater productivity effects if they have a set of intrinsic characteristics capable of boosting economic activity and generating jobs more consistent with the available stock of formal human capital. In addition, they indicate to a great importance of both unobserved individual and regional skills in the formation of more efficient worker-occupation pairs.

4.3 Matching and Reverse Causality

The Column (2) of Table 6 shows the results of a two-stage least square regression on graduate's job matching, where the log of employment density is instrumented by the Sectoral Bartik, without including both individuals and arrangements unobserved heterogeneities.

Table 6 – Impact of Employment Density on Graduates' Job Matching. IV Regressions

	(PLPM)	(2SLS - Bartik)	(2SLS – Population 1910)	(PLPM)	(2SLS-Bartik)	(2SLS-Population 1910)
				With skills	With skills	With skills
	(1)	(2)	(3)	(4)	(5)	(6)
Dens. Empl.	0.00548	0.00374***	0.0027***	-0.0141***	-0.0153***	-0.0189***
	(0.00356)	(0.000136)	(0.0002)	(0.000165)	(0.000488)	(0.000124)
Individuals FE				0.0710***	0.0367***	0.0999***
				(0.00296)	(0.00457)	(0.0002247)
Agglomer.FE				0.0946***	0.0982***	0.0989***
				(0.00793)	(0.0213)	(0.00546)
HHI Industry	0.208	0.103***	0.2620***	0.0494***	0.0628***	0.0122***
	(0.121)	(0.00367)	(0.00651)	(0.00368)	(0.00600)	(0.00153)
HHI Ocuppat.	0.889***	0.353***	0.7483***	0.127***	0.0321*	0.02767***
	(0.270)	(0.00959)	(0.01624)	(0.00962)	(0.0141)	(0.00333)
Individuals	Yes	Yes	Yes	Yes	Yes	Yes
Firms	Yes	Yes	Yes	Yes	Yes	Yes
Geographic	Yes	Yes	Yes	Yes	Yes	Yes
Years	Yes	Yes	Yes	Yes	Yes	Yes
Obs	25943111	25943111	25943109	25983841	25943109	25943109
R^2	0.095	0.0948	0.0920	0.098	0.7832	0.7804
F		2851.23	2772.19	1.6e+08	1.2e+08	1.6e+08
Prob>F	0.000	0.000	0.000	0.000	0.000	0.000
			First Stage Statistics			
Bartik/Pop1910		0.5931	0.4983		0.3297	0.0691
Prob > F		0.0000	0.0000		0.000	0.0000
R^2		0.8630	0.6264		0.9090	0.7804
F (Excluded)		36.62	8.27		92.7	10.01
And-Rubin		42.80	12.89		83.42	31.05
KP Weak-LM		10.07	20.36		26.33	15.423
Correlation	Dens.Emp.	P-value				
Bartik (log)	0.8864	0.0000				
Pop1910 (log)	0.5023	0.0000				
Bartik R^2	0.6773	0.0000				
Population R^2	0.4949	0.0000				

Source: Authors based on RAIS 2010-2014. Note: Standard errors are clustered at the individual level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The coefficient of the impact of density on graduates' job matching dropped 31% in comparison to the PLPM model (1), suggesting that endogeneity has a relevant impact on the identification of the model. As in the model with Bartik IV, the estimate of the impact of employment density on the match, with the population log in 1910 as the instrumental variable, dropped substantially. This finding corroborates the endogeneity between match and urban density and suggests a greater diversity of productive opportunities in agglomerated environments, fostering better qualification-employment matches, while such environments should attract more qualified workers. However, compared to model (2), the estimate was halved.

The strong correlation between the sectoral Bartik and the density of employment shows that this variable estimates a very close value for the distribution of employment in the Brazilian agglomerations. At the same time, the large R^2 (67.73%) and significant at 1% suggests that variations in regional employment are largely driven by changes in the structure of national employment.

By including unobserved characteristics of individuals and agglomerations (columns 3 and 4), the results show a change in the density effect. The relationship turned to be negative in both models. Specification with instrumental variable differs marginally from the POLS model with controls for unobservable attributes of individuals and regions (-1,41% and -1,53%, respectively). Thus, job matching endogeneity in Brazil has roots, fundamentally, in the sorting of skills.

The negative coefficient, even after identification, reveals, however, that the channels of transmission of productivity effects from the agglomeration do not pass-through matching. The effects of urban density on the likelihood and quality of matches disappear when individual skills and regional attributes are controlled. On the other hand, the sign change, at the same level of significance, confirms the idea that the urban density promotes competition among workers. The selection criterion, therefore, becomes the capacity for more productive use of formal qualification, which, in large part, is associated with the stock of individual unobservable skills. At the same time, job matching seems to be related with the nature of the local productive specialization and the disposition of the skills in regional scope. In short, the graduates' job matching in Brazil apparently derives from a sectorial / occupational structure that is spatially specialized and organized around the concentration of skills. In this sense, environments that concentrate more skilled workers must specialize in more skilled economic activities, fostering more complex jobs and attracting more qualified individuals.

5. Matching and Urban Wage Premium in Brazil

Several studies in Brazil have pointed to an urban wage premium, even controlling for potential sources of endogeneity (Rocha et al., 2011, Barufi et al., 2016, Silva, 2017). However, indications that higher labor productivity in large urban centers reflects both the concentration of more skilled workers and the better matching require the identification of matching effects on the urban wage premium. Figure 4 shows that density and matching positively influence the urban wage premium (GLAESER; MARÉ, 2001, ANDERSON *et al*., 2007, COMBES; GOBILLON,2008). The wages, however, seem to be influenced more strongly by variation in matching. In terms of elasticity, the 1% increase in the amount of job matches leads to a 4.7% increase in wages. In turn, the 1% increase in employment density leads to a wage increase of up to 2.1%. In this sense, matching and urban density seem to be potentially explanatory factors for the urban wage premium.

On the other hand, the endogeneity between matching and employment density seems to lead to confusion between the effects of both on worker productivity. In this case, endogeneity would lead to misinterpretations of the impact of the labor market size on the urban

wage premium. In practice, both sorting and matching can inflate these estimates and not handling them can lead to biased estimates of urban density effects on wages. Likewise, not controlling for the effects of urban density can lead to biased estimates of the contribution of the matching effect on wages.

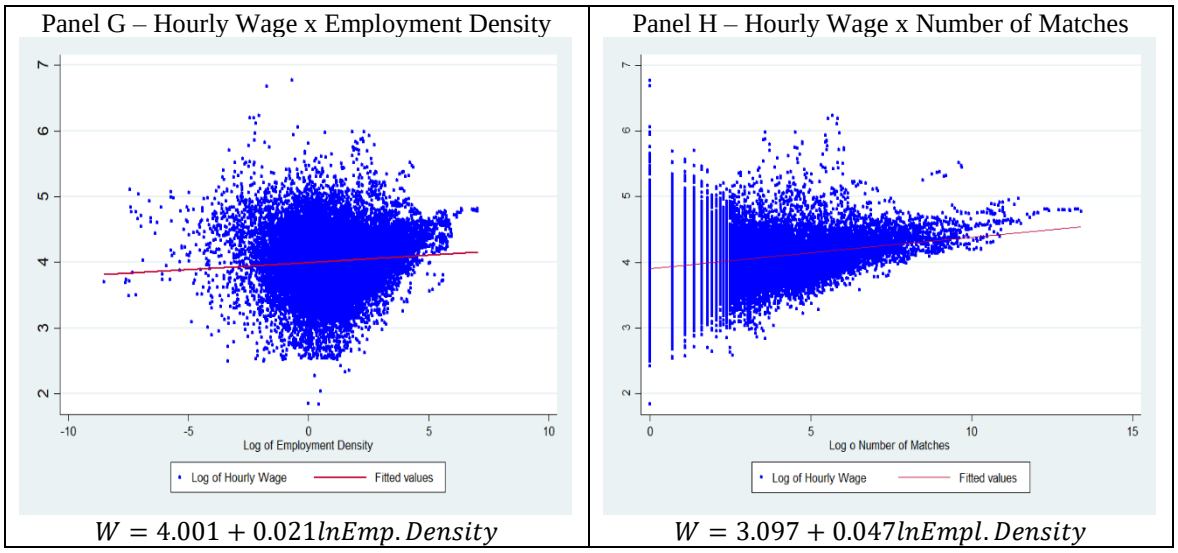

Figure 4 Hourly-Wage, Employment Density and Matching in Brazilian Agglomerations
Source: Authors based on RAIS 2010-2014.

Thus, we investigate the importance of job matching as agglomeration economies driven. To do so, we estimate the following models:

$$lnw_{ikt} = \alpha + \beta EmpDens_{j,t} + \gamma X_{ijt} + \delta F_{jt} + \mu Z_{jt} + \emptyset_i + \sigma_j + \varepsilon_{ijt} \quad (4)$$

$$lnw_{ikt} = \alpha + \beta' EmpDens_{j,t} + \Gamma Match_{ijt} + \gamma X_{ijt} + \delta F_{jt} + \mu Z_{jt} + \emptyset_i + \sigma_j + \varepsilon_{ijt}$$

Where $Match_{ijt}$ is the matching of individual i, in agglomeration j. Γ is the contribution of matching to the wage premium and β' represents the urban wage premium arising from all other sources of urban agglomeration, except matching. In these terms, the contribution of matching to urban productivity can be measured by the difference between β and β'.

Most studies on urban economics for Brazil report matching as a potential source of agglomeration economies. At the same time, they strongly suggest that the productivity effects associated with denser environments derive from better employment matching in favor of larger cities. However, there are no studies that identify in an isolated way the effects of employment matching on urban productivity. This work, in this sense, additionally contributes to the literature by isolating the influence of job matching on the wage premium. In other words, to the extent that individual, firm and region attributes are controlled, the proposed model allows the disaggregation of the wage premium into a component related to concentration (urban wage premium) and into a wage component related to the best structure of job matches (urban matching premium). To control for the sorting and unobserved regional attributes we also include individual and agglomerations fixed effects ($\emptyset_i$ and σ_j). In order to control reverse causality, we estimate the model by using instrumental variables. In this way, it is possible to correct the reverse causality related to the fact that individuals are attracted to places with high wages. At the same time, it is possible to identify the extent to which job matching in denser environments influences urban productivity.

The first three columns of Table 7 show the coefficients of the PLPM models, with controls gradually included. Models (1) and (2) show that both density and matching positively

impact on the graduate's productivity. The positive coefficient of employment density reflects an urban wage premium. However, Column (3) suggests that the match premium is 6 times higher than the density premium, in line with findings of Abel and Deitz (2014). This pattern of results indicates that better matching fundamentally explains the differences in productivity between qualified workers.

The change in the employment density coefficient after controlling for matches was only marginal. Evidence suggests that matching accounts for approximately 3% of the urban wage premium. Similarly, the urban density coefficient is not significantly affected when matching is excluded of regressions (column 1). On the contrary, there is a certain stability, suggesting that the effects of urban density on wages are not driven by the better matching in denser environments. The positive relationship, in this sense, reveals that urban density plays a fundamental role in determining the wage premium, however it operates through channels that do not necessarily involve job matching. In column (4) we include an interaction term between density and job matching. The coefficient was not statistically significant. The employment density coefficient, in turn, remained positive and highly significant. The effect of the match on wages also did not vary, remaining positive and significant at 1%. Evidences, in this sense, points out that there are no systematic differences in the size of the matching premium between large and small cities.

Table 7 – Impact of Employment Density and *Matching* on wages of Brazilian workers

	PLPM	PLPM	PLPM	PLPM	Individual FE	Individual, Agglom. FE	Individual, Agglom.FE Bartik IV	Individual, Agglom.FE Pop1910 IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(7)
DensEmp	0.078***		0.0758***	0.0773***	0.0483***	0.0480	0.0246*	0.0231***
(Log)	(0.0002)		(0.00018)	(0.00021)	(0.00017)	(0.0389)	(0.0113)	(0.0113)
Match		0.443***	0.440***	0.447***	0.168***	0.0613***	0.0645***	0.0659***
		(0.0005)	(0.0005)	(0.00067)	(0.0003)	(0.00817)	(0.00027)	(0.00027)
Match x				-0.00001				
Dens.				(0.0100)				
Individual	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firms	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industries	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geograph.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Years	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	28576647	28576647	28576647	28576647	28576626	23919733	23215981	22635723
R^2	0.355	0.391	0.404	0.404	0.3575	0.902	0.199	0.3905
(Prob)>F	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
First Stage Statistics								
Bartik							0.1327***	0.0690***
							(0.00001)	(0.0004)
R^2							0.8546	
F							2200000	1900000
Prob>F							0.0000	0.0000

Source: Authors based on RAIS 2006-2014. Note: Standard errors are clustered at the individual level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

When we control for fixed effects (column 5), matching coefficient dropped to 16%, while density turns out to explain only 4% of the wage premium. Following the empirical evidence already found for Brazil (Rocha et al., 2011, Barufi *et al*., 2016, Silva, 2017), the results point that sorting of skills is the greatest source of endogeneity in wages. On the other hand, matching seems to be intrinsically related to the regional concentration of skills. Skilled markets are more complex, concentrate more technology-intensive occupations, and promote tighter interactions between firms and skilled workers.

Column (6) reports to the model with simultaneous fixed effects for individuals and regions. The coefficient of employment density did not systematically change, but it was not statistically significant. The match premium estimate, however, is robust, but it dropped to 0.0613 (it fell about 63%), reinforcing both the idea that there are no systematic interactions between urban density and matching and that productivity effects of matching are strongly correlated with sorting and regional attributes. On the other hand, the urban wage premium is not drawn from the local job matching structure, but from the regional ability to generate productive opportunities for workers – condition intrinsically related to the concentration of skills.

As individuals are attracted to locations with higher wages, reverse causality can influence the effect of job density and matching on the wage premium. Column (7) addresses the simultaneity between wages and density, using the variable Bartik as an instrument for job density. The reported statistics reiterate the strength and relevance of the instrument. The second stage estimates show that both the match of skills and the density of employment generate wage premiums. Even after controlling for reverse causality, the match premium remained stable. Matching seems to be driven primarily by skills and the channels of transmission of their effects on productivity depend especially on how workers make use of them. On the other hand, the density coefficient dropped by approximately 50%, with a significance of 5%. The urban wage premium estimate is similar to that found in Silva (2017), around 2.6%, and suggests agglomeration economies acting on wages. However, the substantial decrease in the size of the estimated effect of density suggests that the effects of agglomeration reflect the endogenous nature of the relationship between wages and density. The sorting of skills, however, seems to concentrate an important part of this endogeneity.

In short, the results indicate that both matching and density matter in explaining differences in productivity. Individuals with better qualification and occupation matches experience higher wage gains. The same is true among workers in denser labor market areas. The comparison between the different models, however, confirms the prediction that the influence of urban density on wages is not driven by better matching of skills in more crowded environments.

5. Conclusion

The paper investigated the relationship between urban agglomeration and job matching for college graduates in Brazil. The research contributes by addressing the causal relationship between the labor market size and the education-occupation matching, considering different agglomerations in a large and heterogeneous geographical area. We use a panel data of college graduated workers in the period from 2006 to 2014 to investigate the influence of sorting and unobserved regional attributes on matching into more urbanized areas.

We conclude that workers in denser markets are more likely to develop better occupation and education matches. That is, the main evidences show that denser markets leverage better worker-job pairs. The impact on matching is robust to an extensive set of observable characteristics of workers and region. Overall, even after these controls, an 1% change in employment density increases the probability of matching by up to 0.25 percentage point. This result is consistent with the findings of Abel and Deitsz (2014). However, in order to deal with sorting, we take advantage of the panel structure to estimate individual and region fixed-effect models. When such models are applied, the estimates turn out to be negative, suggesting that influence of urban density on matching is inflated by the sorting of skills and regional attributes. We also address potential reverse causality by instrumenting current

employment density with Bartik IV and also with historical population data. Results are very similar to baseline and fixed effects and corroborate that the endogeneity between matching and agglomeration is fundamentally based on sorting of skills. Thus, differences in the regional skill composition seem to account for the great part of the matching differential across regions in Brazil.

It is worth mentioning that the coefficients of occupational and sectorial specialization, elucidated by the Hirschmann indexes, proved to be strongly related to the probability of matching. Thus, evidence suggests that, in the Brazilian case, specialization is more important than location in explaining the correspondence between employment and education. It makes sense, if we analyze the regional labor markets from the perspective of the sectoral organization of economic activity, whereby the Center-South regions concentrates more sophisticated establishments, while the North-Northeast still has economic activities based on basic services and industries of low economic complexity. In Brazil, which is a country with marked regional heterogeneities, the way in which productive activity is organized seems to be a relevant factor in explaining regional inequalities in opportunities.

Finally, we estimate wage equations in order to identify whether there are systematic interactions between the urban wage premium and the matching premium. Estimates showed that both the match of skills and the density of employment generate wage premiums. However, there seems to be no systematic differences in the size of the matching premium between small and large cities. The inclusion of unobserved components achieved substantially lower estimates for both the match and the employment density, suggesting that the concentration of individual skills constitutes a fundamental factor in explaining the wage premium. In these terms, there are strong indications that matching is driven primarily by skills and that the channels of transmission of their effects on productivity depend especially on how workers use them.

The results presented here reveal only one dimension of job matching – the matching between formal qualifications and job requirements. The limitations imposed by the database allowed us to focus only on the concept of vertical matching, through which it is possible to identify the amount of education needed to perform certain tasks. However, it is important to emphasize that skill matching is a concept with several dimensions of analysis and of greater amplitude that can help to understand the differences in productivity between locations (Abel; Deitsz, 2014, Berlingieri *et al*., 2016). In these terms, future studies may consider the contribution of different matching measures as sources of agglomeration economies. Furthermore, the inclusion of firm effects may provide more accurate evidence regarding the formation of good worker-job pairs. Another possible extension is whether migration to denser regions affects, and to what extent, the probability of matching. Finally, the limitations of this research allowed adding evidence only on the probability of matching. The effects of agglomeration on its quality involve the construction of horizontal matching measures, so that it is possible to relate the specific field of higher education to the job performed by the individual.

References

ABEL, J.; DEITZ, R. Agglomeration and job matching among college graduates. **Regional Science and Urban Economics**, [s.l.], v.51, p.14-24, 2014.

ANDINI, M.; DE BLASIO, G.; DURANTON, G.; STRANGE, W. C. Marshallian labour Market pooling: evidence from Italy**. Reg. Science and Urban Economics**,v.43, p. 1008–1022, 2013.

BARTIK, T. 1991. **Who Benefits from State and Local Economic Development Policies?** Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 1991.373p.

BARUFI, A.M., HADDAD, E.A. Spatial Mismatch, Wages and Unemployment in Brazilian Metropolitan Areas. **The Journal of ERSA**, [s.l.], v.4, n.3, p.175-200, 2018.
AMARANTE, P.; BATISTA DA SILVA, M. V.; MONTE, P.A. Does the spatial density of employment stimulate inter-firm worker mobility? An analysis of Brazilian municipalities. **The Annals of Regional Science**, [s.l.], v. online, p. 1-31, 2019.
BERLIANT, M.; KONISHI, H. The endogenous formation of a city: population agglomeration and marketplaces in a location-specific production economy. **Regional Science and Urban Economics**, [s.l.], v. 30, n.3, p. 289-324, 2000.
BERLINGIERI, F. Local Labor Market Size and Qualification Mismatch. **Journal of Economic Geography**, Oxford, p.1-26, 2018.
BERRY, C. R.; *GLAESER*, E. L. The divergence of human capital levels across cities. **Papers in Regional Science**, [s.l.], vol 84, n.3, 2005.
BOUALAM, B. Getting a First Job: Quality of the Labor Matching in French Cities. **Working Paper Series**, Genebra, 2014 (mimeo), 53p.
COMBES, P.; DURANTON, G.; GOBILLON, L. Spatial wage disparities: sorting matters! **Journal of Urban Economics**, [s.l.], v.63, n.2, p.723–742, 2008.
COMBES, P.-P.; DURANTON, G.; GOBILLON, L.; ROUX, S. Estimating agglomeration economies with history, geology, and worker effects. In: Glaeser, E.L. (Org.), **Agglomeration Economics**. Chicago: University of Chicago Press for the **NBER**, 2010. p. 15–65.
D'COSTA, S.; OVERMAN, H. G. The urban wage growth premium: Sorting or learning? **Regional Science and Urban Economics**, [s.l.], v.48, p.168-179, 2013.
DURANTON, G.; PUGA, D. Micro-foundations of urban agglomeration economies. In: HENDERSON, V.; THISSE, J.F. (Orgs). **Handbook of Regional and Urban Economics**, [s.l.]: Elsevier, v. 4, 2004. p.2063-2117.
GLAESER, E.; MARÉ, D. Cities and Skills. **Journal of Labor Economics**, Chicago, v. 19, n.2, p. 316-342, 2001.
HELSLEY, R. W.; STRANGE, W. C., Matching and agglomeration economies in a system of cities, **Regional Science and Urban Economics**, [s.l.], v.20, n.2, p.189-212, 1990.
KORPI, Martin; CLARK, William AV. Migration and occupational careers: The static and dynamic urban wage premium by education and city size. **Papers in Regional Science**, v. 98, n. 1, p. 555-574, 2019.
KRUGMAN, P. Increasing Returns and Economic Geography. **Journal of Political Economy**, [s.l.], v.99, n.3, p.483-499, 1991.
MACIENTE, A.N. **The determinants of agglomeration in Brazil: input-output, labor and knowledge externalities**. PhD Dissertation – Univ. of Illinois at Urbana-Champaign, 2012.
OVERMAN, H. G.; PUGA, D. Labor Pooling as a Source of Agglomeration: An Empirical Investigation. In: GLAESER, E.L. (Org.). **Agglomeration Economics**. New York: NBER Chapters, 2010. p.133-150.
RAUCH, J.E. Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities. **Journal of Urban Economics**, [s.l.], v.34, p.380-400, 1993.
ROBST, J. Education and job match: The relatedness of college major and work, **Economics of Education Review**, [s.l.], v. 26, n.4, p.397-407, 2007.
ROCHA, R.M; SILVEIRA NETO, R.M; GOMES, S.M.F.P.O; Maiores Cidades, Maiores Habilidades Produtivas: Ganhos de Aglomeração ou Atração de Habilidosos? Uma Análise para as Cidades Brasileiras. In: **Anais**, ANPEC NORDESTE, 16, Fortaleza, 2011.
SHIMER, R.; SMITH, L. Assortative matching and search. **Econometrica**, [s.l.], v.68, n.2, p. 343–369, 2000.
SILVA, D.G. **Economias de Aglomeração e Heterogeneidade de Trabalhadores e Firmas na determinação de salários no Brasil**. 2017. 160f. Tese (Doutorado em Economia) – FEA-USP, São Paulo, 2017.