

Skill wage premium and city size: evidence from Brazil

Edivaldo Neves Jr., Carlos Azzoni, André Chagas
NEREUS, Department of Economics, University of São Paulo

Abstract

We estimate the urban wage premium for the Brazilian case, exploring how it is heterogeneous for workers/occupations with distinct skills. Every worker/occupation was assigned a level of cognitive, social and motor skills. By using a rich data database matching firms and workers, we replicate the wage premium results of other studies. However, we were able to show that the wage premium does not affect equally all occupations/workers. We found a positive association between cognitive and social skills and urban size, especially for workers/occupations that use intensively those skills, and a less intense or non-significant effect for motor skills. Returns to cognitive skills are positive across the whole range of urban sizes, but returns to social skills are only present in large urban agglomerations. Furthermore, interactions of social and cognitive skills inflate the skill wage premium associated with urban size.

Keywords: Urban premium; skills; sorting;
JEL: R23; J24

1. Introduction

The evidence on the relationship between wage levels and urban size is well established (Glaeser and Mare, 2001; Glaeser and Resseger, 2010; Combes *et al.*, 2010, Ronsenthal and Strange, 2004; Overman and Puga, 2010; Puga, 2010). Explanations are plenty, such as the Marshallian triad or the micro-foundations models developed by Duranton and Puga (2004), exploring topics as *sharing*, *matching*, and *learning*. Larger agglomerations increase the opportunities for sharing infrastructure, suppliers and a labor pool with similar skills, amplify the chances of matching between workers and firms in the labor market, and create a higher probability of learning and developing new technologies and entrepreneurial practices (Puga, 2010; Combes *et al.*, 2010). Regional wage differences could also come from different labor force compositions, geographic characteristics and local production factors, and agglomeration economies (Combes *et al.*, 2008).

Endogeneity plays an important role in this discussion: Do the city characteristics make economic agents more productive, or are large cities more productive because they have more productive economic agents? There could be non-observable individual characteristics correlated with urban size biasing the identification of the impact of size on wages (sorting). That is the case if skilled workers prefer urban centers with better social, economic or urbanist attributes, as well as the previous existence of skilled people (Carlsen *et al.*, 2012). Therefore, considering non-observed worker characteristics in the calculations is extremely relevant, as did Glaeser and Mare (2001), D'Costa and Overman (2014), Mion and Naticchioni (2009).

Besides investigating the sources of the urban wage premium, it is important to consider how the urban wage premium accrues to workers with different skill profiles. The idea is to find what types of individuals play a key role in the increase in urban productivity. One way of differentiating individuals is through education levels, but this is highly unsatisfactory. The type of activity performed seems to be more relevant than the educational level (Bacolod *et al.*, 2009; Florida *et al.*, 2011; Maciente, 2013). In this study, we investigate the relationship between productivity (wage) and urban size for occupations (workers) with different skill requirements. We study the spatial distribution of individual skills and estimate the impact of agglomeration on the hedonic price of skills.

This approach is still in its infancy, especially in the literature in Brazilian cities. We use a database of administrative records for firms and workers, including the occupation of each worker. The occupations were assigned skill levels, based on an adaptation of the American ONET survey to the Brazilian labor market (Maciente, 2013). For each occupation/worker, we have a skill score for cognitive, social and motor skills. The relative frequency of individuals with high cognitive and social skills increases as the urban population increases, whereas for motor skills the opposite is observed. The econometric exercises show that the importance of cognitive and social skills grows as the city size increases, whereas for motor skills it happens with less intensity. The following sections present a brief discussion of the literature and introduce the theoretical model. Following, we present a description of the database and the construction of the skill indicators used. The empirical strategy is presented and the results are revealed next, together with the conclusions of the study.

2. Literature

In order to discuss the relationship between skills, productivity and urban size it is important to reflect on the meaning of agglomeration economies. Combes and Gobillon (2015) argue that they cover any effect that increases the income of workers and firms as the city size increases. Groot and De Groot (2014) consider that they are wage differences resulting, *ceteris paribus*, from the proximity of different firms and consumers, dense labor markets, and knowledge spillovers. Glaeser and Gottlieb (2009) link economies of agglomeration with the reduction in the cost of transportation of goods (proximity between suppliers and buyers of goods and services), people (labor market is more efficient in urban areas), and ideas (cities facilitate the flow of knowledge between firms and persons).

From an empirical point of view, identifying the channels leading to the existence of agglomeration economies is a difficult task (Combes and Gobillon, 2015). However, important advances have been made in the empirical identification of the sorting component. Sorting could occur if individuals with higher skills attribute greater value to urban amenities (culture, institutions, nightlife, etc.) or if, historically, skilled people located in large cities and transmitted the skills to future generations. Ignoring the sorting mechanism will tend to overestimate the impact of agglomeration economies. Empirically, it is necessary to control for individual heterogeneity (Glaeser and Mare, 2001; Combes *et al.*, 2008). The literature indicates that sorting is a relevant source of the urban wage premium (Combes *et al.*, 2010; Carlsen *et al.*, 2010; Matano and Naticchioni, 2015).

However, the evidence on the different magnitudes of the urban wage premium for workers with different skill profiles is scarce. As shown by Bacolod *et al.* (2009), only part of the workers/occupations benefit from the urban size, depending on their skills, and the premium is not uniform for workers with different skills. They indicate that large cities are particularly good for workers/occupations demanding high levels of cognitive and social skills. Florida *et al.* (2011) show that the effects of analytical and social skills on wage are positive, while the effects of physical skills are negative. Explanations include considering that skilled workers tend to specialize and, therefore, benefit from better matching conditions in large cities; this type of workers are able to learn from the rich environment of large cities; they can take advantage from the complementarity of resources available in large and dense markets. Similar conclusions were also present in a study developed by Andersson *et al.* (2014), who found signs of agglomeration economies only for non-routine tasks.

This new perspective of measuring human capital makes it possible to deepen previous studies, offering new insights to analyze issues such as labor market characteristics or the dynamics of cities, providing more subsidies for policy-makers' decisions. The main criticism regarding the use of education to represent skills is that it may not measure properly the degree of qualification of workers, because it does not capture other specific skills that are often developed only in the work environment. For example, a person with 15 years of schooling could have studied Engineering, Physics, Sociology or any other field, but would be considered with the same skill if only schooling were considered. It seems clear that professionals in these fields may have distinct characteristics and skills, and these would be lost if only education levels were considered. Thus, with this approach, it becomes clear that education and skills are not equivalent¹.

The fundamental hypothesis that makes it possible to infer the skill level of the worker from his/her occupation is the existence of matching in the labor market. It is considered that in a frictionless hedonic equilibrium, the labor market associates each worker i , who possesses certain abilities s , for occupations j that require such characteristics, making it possible to map the relation occupation-skill of the worker. If this relation is invertible, it can be argued that the worker assigned to an occupation j has the level of skills necessary to occupy such job.

The Brazilian case is interesting for the diversity of its urban network and the rapid expansion of the urban system in the last six decades. Although a reduction in wage inequality can be observed in recent years (Cruz and Naticchioni, 2012), the evidence indicates a positive relationship between wage levels and urban size, even after controlling for cost of living and characteristics of workers and firms (Azzoni and Servo, 2002). Rocha *et al.* (2011) registered a 9.4% wage premium for metropolitan areas, even after controlling for observable and non-observable characteristics of workers. Freguglia *et al.* (2007) indicated that the measured regional and sectoral wage differences are smaller when non-observable worker characteristics are considered. Silva *et al.* (2012) obtained an urban wage premium of 3%, after controlling for the non-observable individual heterogeneity and the observed characteristics of workers and jobs. None of those studies, however, considered skill diversity. Exceptions are Barufi *et al.* (2016), who used skills as a control. They concluded that, controlling for worker'

¹ Several authors emphasize that skills are multidimensional and not simply identical to educational level (Levy and Murnane, 2003).

skills and urban size, the sectoral composition is relevant. Ehrl and Monasterio (2016) discuss the concentration of analytical and interpersonal skills and their effects on wages, using as instrument differences in the distribution of industrial and liberal occupations in 1872 and 1920. The authors found that the regional concentration of these occupations has positive externality on wages. Finally, Andrade *et al.* (2014) indicated that the occupational concentration is affected by the geographical distribution of productive activities and by the technological intensity of each occupation. As far as we know, there is no study on Brazil dealing with the importance of urban size on the implicit price of skills.

3. The model

We use an adaptation of Roback (1982) model to describe the spatial equilibrium in the presence of agglomeration effects and worker's skill heterogeneity, as in Moretti (2004), Rosenthal and Strange (2008), Bacolod *et al.* (2009) and Liu (2016). Suppose the existence of several locations j , composed of I workers (identified by i). Furthermore, each worker i presents a vector of skills z_i . The welfare of residents in each city is given by the function

$$U=(X,h_j,A_j)$$

In which X is a composed tradable good, whose price is fixed inter-regionally and used as numeraire; h refers to the non-tradable good, typically housing; A is the local agglomeration effect.

Admitting that workers supply one unit of labor, the problem of an individual agent is to maximize his/her utility subject to a budget restriction described by

$$W_j(z_i)+K=X+r_jh_j$$

Given the agglomeration level in the city, A_j , and the prices of goods, wages (W) and rent (r), the individual must choose the quantities of services to be consumed, following the restriction, which is given by his/her income from wages and other sources, K . Given the possibility of locational arbitrage, the equilibrium condition for the worker can be represented in terms of indirect utility as

$$V[W(z_i), r, A]=V^*(z)$$

As for the firms, suppose they operate with constant returns to scale and show the following production function

$$X=f[Nz_i,h_p,A_j]$$

In which Nz_i is the number of workers employed, described by the vector of skills z_i ; h_p indicates the use of land in production, and A_j indicates the agglomeration level in city j . Assume there is free mobility of labor (Nz_i) between localities and local services and housing are fixed for each city. The profit maximization problem is given by

$$Max \pi=X - r_jh_j - W(z_i)Nz_i$$

The optimum wage level resulting from this maximization is

$$W(z_i) = f'[Nz_i,h,A_j]$$

where f' is the derivative of the production function in relation to Nz_i , that is, the marginal product of labor.

Suppose that there are S skills, indexed by s . In this setup, the **skill wage premium**, or the hedonic price of skills, or the return to skills, is given by the marginal contribution to wage of the particular skill

$$(\partial W(z_i)/\partial N_{z_i_s})$$

This is the implicit price of a specific skill. This derivative allows to infer about wage behavior as a function of individual skills. The evidence in the literature indicates that higher skill levels are associated to higher wage levels. This article takes a different point of view, casting a new insight on the relationship between the wage premium and individual skill levels.

The effect of agglomeration on wages, or the **urban wage premium**, is captured by the marginal contribution of land used in production on wage

$$\partial W(z_i)/\partial A.$$

This derivative reveals how important the urban environment is in generating wage returns. The analysis of this derivative received great attention in the identification and estimation of the agglomerative effects associated with city size. However, the presence of unobservable characteristics related to the individual wages present in the error term brings empirical difficulties, since it causes an omitted variable bias problem in the OLS estimator. Even considering this limitation, there is plenty of evidence that wages growth with urban size.

Finally, the **urban-size skill wage premium** is given by the effect of agglomeration on the skill wage premium, that is

$$\partial(\partial W(z_i)/\partial N_{z_i_s})/\partial A = \partial^2 W(z_i)/\partial N_{z_i_s} \partial A.$$

This effect, which is the main interest of this paper, indicates how the implicit price of an individual skill is affected by agglomeration, or how the hedonic prices of skills vary with city size. It allows us to explore the relationship between the increase in population size and the wage return associated with each specific skill. The micro foundation models proposed by Duranton and Puga (2004) indicate the channels through which differences in the relationship between skill premium and urban size are expected. Individuals with a high level of cognitive or social skills are more able to benefit from *learning* in large centers. They are also in a better position to benefit from the *matching* opportunities offered by thicker urban labor markets, or from the *sharing* of the complementary resources present in large cities.

Using the same argument for motor skills is not straightforward, however. Although individuals with motor skills could benefit from *matching* or *sharing*, it is less likely that they could gain from *learning*. Thus, for cognitive and social occupations, there are more channels to justify the existence of a positive relation between the wage premium and city size, whereas for the motor group the argument is weaker.

To close our model, in equilibrium, the problem of the firm can be represented by

$$\pi[(W(z_i), r, A)] = 0.$$

This equation, together with the indirect utility function, determines the different levels of rent and wage as a function of skills, for each place. Then, equilibrium is described by the wage and rent levels in which (a) workers are indifferent between living at place j or any other, and (b) firms obtain the higher profit level.

4. Data and Descriptive analysis

4.1 Data

We use yearly administrative records of firms and workers from a database assembled by the Minister of Labor, covering all firms legally established and workers with a formal labor contract. It contains reliable² information on several socioeconomic and job variables, including the description of the worker's occupation, and firm characteristics (sector, size, location, etc.). The data covers the whole country in the period 2003-2013. As the database includes an intractably large number of jobs, for the panel data analysis we have drawn a random sample of 3% of workers, weighted by municipalities (totaling over 600 thousand workers). For the cross-section analysis we used a 20% sample, for the year 2013, involving close to 6 million workers. Additionally, we restrict our analysis only to the private sector, to prevent specificity of Brazilian public sector.

Each of the 2,702 occupations receives a code based on the international classification of occupations. This allowed Maciente (2013) to produce an adaptation of the American ONET list of skills to the Brazilian scene. In order to reduce the set of skills to a manageable level, we have defined three types of skills to use in this study: cognitive, social and motor. From the list of 263 skills associated with each occupation in the ONET, we have selected three sub-sets corresponding to skills, in principle, more relevant for each of the three categories chosen. We then applied Factor Analysis to each subset in order to reduce the number of variables, without compromising the amount of information present in the original dataset. The three resulting skill indicators are not orthogonal, but preserve a desirable relation of complementarity (Bacolod and Blum, 2010). It is important to note that the choice of skills followed a textual analysis and the evidence in the literature (Ingram e Neumann, 2005; Bacolod *et al.*, 2009; Florida *et al.*, 2011; Weinstein e Partridge, 2013; Guvenen *et al.*, 2015).

Cognitive skills are those related to logical reasoning, to learning capacity and to mastering language; social skills are related to interpersonal relations in the work environment; motor skills are related to manual dexterity and strength to develop physically demanding jobs. A total of 22 variables were selected, 9 for cognitive, 7 for social and 6 for motor skills³ (see Table 2, below). Thus, each worker, based on his/her occupation, receives three values corresponding to the three skills. The correlation between cognitive and social skills is positive and between these two and motor skills is negative. Higher levels of cognitive and social skills are associated with higher wages; higher levels of motor skills, with lower wages.

The urban agglomerations used in this study are 369 Labor Market Areas (LMA)⁴, encompassing 1,939 municipalities, which account for 71% of the population in 2013.

² We have eliminated cases with inactive contracts, zero-wages, less than 20 hours/week, public workers, ages below 18 and over 65, and some clearly defective data (such as different gender in distinct years, decreasing age, etc.).

³ Other ways of constructing the skill variables were tested. We applied Factor Analysis to the whole set of skills. We rotated the factors obtained using the PROMAX technique, to maintain the complementarity between the skills. Alternatively, based on the explanatory power of the skills variables present in Maciente (2013), new skills indexes were created. The key results of this study remained.

⁴ The LMAs include the 294 officially established by IBGE, the national statistics office, based on commuting to work and study, plus 75 large cities, including the municipalities under their influence in the urban hierarchy (satellites), as defined by IBGE. In order to check for robustness, other definitions

Few LMAs involve several municipalities, mostly concentrated in the richer south and southeast regions, and 76% are composed of five cities or less. To facilitate the analysis, we have defined a top category for of each skill, composed of workers situated among the 20% occupations with the largest scores, and four sizes for the LMAs. Table 1 provides some statistics, and shows that 50% of individuals work in very large LMAs, 60.4% for the top-cognitive workers.

Table 1 – Skills and wages in LMAs (2013)

	All workers	Top-Motor	Top-Cognitive	Top-Social
% in large LMAs (>2.5 million)	50.2	43.9	60.4	51.66
Average wage	2,078	1,626	5,736	3,185
Large LMAs	2,345	1,663	6,563	3,916
Small LMAs (< 100,000)	1,567	1,497	3,508	1,904
% Difference	33.1	9.94	46.5	51.37

Source: Organized by the authors.

The average wage of these top-cognitive workers is over twice the general wage average, and the difference is larger in large LMAs. The group of top-social workers presents an average wage lower than the top-cognitive, but higher than the general average. Workers in the top-motor group get lower wages and are typically less concentrated in large LMAs. Wage levels in large LMAs for the top-cognitive occupations are 46.5% higher than in small LMAs, and only 9.94% for the top-motor skills. The larger return occurs for the top-social skills, for which working in a large LMA results in doubling the wage level, as compared to working in a small LMA.

Table 2 presents information on schooling and skill levels across urban sizes. Skills are presented in the synthetic indicators constructed with Factor Analysis and for each individual skill considered. It is clear that the share of people with more education increases with urban size. The share of people with a college degree doubles from small to large cities. Regarding skills, 2.61% of individuals working in small urban concentrations have top-cognitive skills, while in very large agglomerations this percentage is 5.7%, and this share increases monotonically with urban size. The pattern is repeated for each individual skill within this group. In the social skills, growth with city size is also repeated, although the difference between small and large cities is not as large. Again, the pattern is present for each individual skill within that group. Finally, the motor skills present the opposite pattern, decreasing as the city size increases, both for the synthetic indicator and for each individual skill in the group.

These results contrast with the findings of Bacolod *et al.* (2009) for the USA, where the average indicators of skills and education are homogeneous across city sizes. Eeckhout *et al.* (2010) argue that the skill distribution presents a *fat tail*, and show evidence that more skilled individuals locate disproportionately more in large cities, and, at the same time, large cities attract less skilled people too. Our results indicate that the distribution of skills is not uniform across Brazilian cities, but no conclusion can be drawn about the existence of a *fat tail* phenomenon.

were also used: a) the 294 official LMAs, and b) these plus the 75 large cities, without the satellite cities. The main results of the study do not change.

Although skill and education are different dimensions of human capital, they are surely correlated. Within workers with a college degree, 30% are in top-cognitive and 38.9% in top-social occupations. Only 0.2% of workers with 8 years of schooling are in top-cognitive occupations, and 33% in top-motor occupations.

Table 2 – Education and skills across urban sizes (2013)

	Urban Size (1,000)			
	<100	100 - 750	750 – 2,500	> 2,500
Schooling (years)				
< 8	41.51	37.42	32.61	30.31
8 – 11	51.10	53.29	56.59	54.84
11 and +	7.39	9.29	10.80	14.84
Skills				
Top-Cognitive	2.61	3.57	4.31	5.70
Deductive Reasoning	3.11	3.98	4.51	6.13
Inductive Reasoning	2.46	3.20	3.81	4.59
Category Flexibility	3.86	5.07	5.39	6.82
Reading Comprehension	1.61	2.14	2.59	3.43
Writing	2.72	3.28	3.87	4.97
Critical Thinking	2.36	3.10	3.77	5.22
Complex Problem Solving	3.16	4.33	5.08	6.85
Analytical Thinking	2.40	3.21	3.94	5.64
Mathematical Reasoning	6.73	8.29	9.09	11.33
Top-Social	16.89	17.04	17.48	18.23
Social Perceptiveness	5.03	6.07	6.88	8.30
Coordination	7.08	8.28	9.27	11.21
Persuasion	7.44	8.51	9.23	10.81
Negotiation	15.10	15.80	17.64	19.15
Establishing and Maintaining Interpersonal Relationships	16.94	16.77	17.20	18.18
Selling or Influencing Others	17.53	18.05	19.77	20.59
Resolving Conflicts and Negotiating with Others	9.48	10.79	12.18	14.40
Top-Motor	23.77	22.23	20.85	16.96
Manual Dexterity	13.84	15.30	15.25	13.29
Control Precision	22.87	18.89	13.26	10.65
Static Strength	31.94	29.09	27.76	23.40
Dynamic Strength	25.22	22.54	22.28	17.98
Performing General Physical Activities	21.19	21.05	21.17	17.67
Handling and Moving Objects	29.07	28.55	24.58	21.59

Source: Organized by the authors.

4.2 Descriptive Analysis

This section aims to provide initial evidence on the relationship between wages, agglomeration and skills in Brazil by exploring the cross-section data structure of 2013. The hedonic skill price equations were estimated with the following econometric specification:

$$\ln w_{istj} = \phi_{st} \ln(\text{pop}) + \beta_{st} Z_{tj} + \mu_{st} Z_{tj}^* \ln(\text{pop}) + \alpha_t X_{ist} + \delta_i Q_{ist} + \varepsilon_{istj} \quad (1)$$

with w_{istj} standing for hourly wages of individual i in occupation j , residing in LMA s , in year t . X_{ist} is a set of worker's characteristics, including race, gender, and schooling. Q_{ist} includes firm size and sector of activity. Z_{tj} indicates the skill level of workers in occupation j , with hedonic prices varying with location s .

The cross-section estimated β_{st} coefficients of Equation (1) show how wages (productivity) vary for workers with different skills - the skill wage premium, or the skill implicit hedonic price. By looking at the μ_{st} coefficients, it is possible to investigate how the skill wage premium varies with city size, by interacting skill levels and urban size – the urban-size skill wage premium. From these results, it is possible to determine urban thresholds for the occurrence of returns to each skill. Besides that, the previous models consider that skills have the same marginal effects on wages, but the marginal contribution can differ for workers with different skill intensity. Thus, we investigate the existence of possible nonlinearities in the specification of the skill wage premium and the urban-size skill premium.

Table 3 – Cross-section Regression Results (2013)

Ln (hourly wage)	1	2	3	4	5	6	7	8
Ln(Pop)	0.035*	0.031*	0.048*	0.032*	0.035*	0.032*	0.032*	0.044*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Cognitive		3.182*	-1.040**					0.194
		(0.052)	(0.437)					(0.577)
Cognitive*Ln(Pop)			0.288*					0.199*
			(0.032)					(0.043)
Social				1.454*	-1.121*			-0.907***
				(0.053)	(0.420)			(0.522)
Social*Ln(Pop)					0.176*			0.067***
					(0.032)			(0.039)
Motor						-1.004*	1.272*	0.478
						(0.053)	(0.436)	(0.371)
Motor*Ln(Pop)							-0.156*	-0.033
							(0.033)	(0.028)
Constant	0.567*	0.971*	0.711*	0.768*	0.705*	0.689*	0.669*	0.765*
	(0.061)	(0.042)	(0.047)	(0.048)	(0.048)	(0.055)	(0.053)	(0.040)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R2	0.4458	0.5154	0.5179	0.4789	0.4807	0.4584	0.4597	0.5181
Number of Obs.	5,986,159	5,986,159	5,986,159	5,986,159	5,986,159	5,986,159	5,986,159	5,986,159
F	711.17	849.70	1443.26	674.65	863.52	687.92	728.67	1327.16
Prob>F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Source: Organized by the authors. Standard errors in brackets are clustered at the occupation/LMA level. *** p<0.1, ** p<0.05, * p<0.01

Table 3 exhibits the estimated skill wage premium. The first model (column1) includes only population and the control variables, and indicates an urban wage premium of

3.5%. The controls presented the expected signs: wage levels grow with firm size and worker's education, and are larger for white, male, and older workers (age squared non-significant). These results are robust and are replicated in all models.

The next columns deal with one skill at a time, and the last column includes all skills simultaneously. In all cases, the controls are included and present the expected signs. The estimated cognitive skills wage premium (column 2) and the urban-size skill wage premium (column 3) are positive and significant. An increase of one standard deviation (0.1) in the cognitive index is associated with a 2.88% increase in the urban-size skill wage premium. This result shows that larger urban sizes are associated with higher wage returns for individuals with cognitive skills.

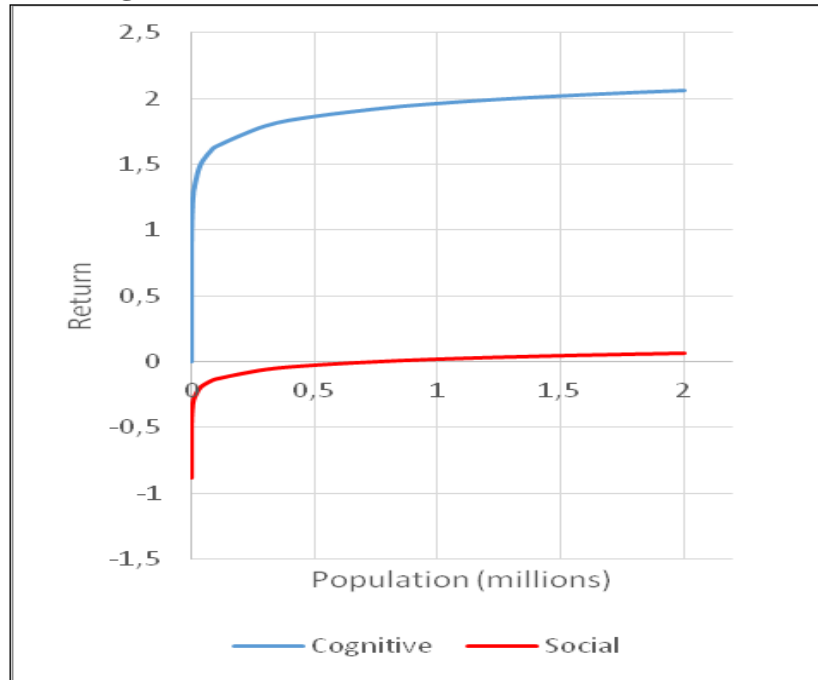
The estimated social skills wage premium is also positive (column 4), but with less intensity in comparison to cognitive skills. The interaction *skill-urban size* in column 5 indicates that an increase of one standard deviation in the social skill indicator is associated with an increase of 1.76% in the urban-size skill wage premium. Columns 6 and 7 indicate that the motor skill wage premium is negative, and decreases with urban size. This set of results indicates that the urban wage premium is directly associated with cognitive and social skills, but not with motor skills.

Finally, model 8 includes all variables simultaneously and the results reinforce the previous conclusions, but the intensity of the effects is in general lower than before, and motor skills become non-significant. As robustness checks, we have estimated similar regressions with the individual skills replacing the synthetic indicators produced by Factor Analysis. The results are replicated for every individual skill, both for the skill wage premium and the urban-size wage premium. We have also estimated interactions between the control variables and population, to check for the possibility that those controls also vary with the population. Again, the results of interest remain. The model described in *Equation (1)* also allows for determining the population threshold from which the returns to the skill show up. The skill wage premium is given by $\partial \ln w_{istj} / \partial Z_{ij}$ and the result of this derivative is $[\beta_{st} + \mu_{st} \ln(\text{pop})]$. Analyzing the sign and finding the root of this expression allows for determining the population threshold. Figure 2 presents the thresholds for cognitive and social skills, based on the results of model 8 in Table 3⁵. The formats of the curves are directly related to the functional form of the equation, to the sign of the coefficient μ_{st} and to the size of the coefficient associated with the skills, β_{st} . The later corresponds to a shift in the return to skills curve. If it is positive, the minimum level of population at which there are returns to the skill is smaller; if it is negative, the population threshold is larger.

It can be seen that occupations intense in cognitive skills find return at any city size, and this return increases with city size. The situation is different for social skills: there are no returns for cities with a population below 757,147 inhabitants. That is, social skills provide positive returns only after a certain city size. Thus, the returns to skills are positive for cognitive skills, for social skills from a certain urban size on, and non-existent for motor skills. It seems, thus, that individuals with cognitive skills are better equipped to absorb the knowledge spillovers existent in urban centers, and that reflects on their wage levels. Individuals with social skills need larger cities to fully benefit from them.

⁵ As the results for motor skills are non-significant, it does not make sense to explore them here.

Figure 2 – Return to skills and urban size (2013)



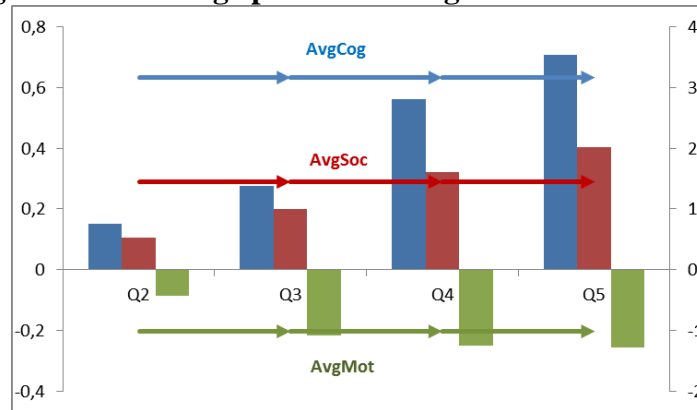
Source: Organized by the authors.

Social skills have received great attention in the literature recently. One of the ideas underlying this growing interest in social skills is that evidence suggests that technological advancement has an important effect on the labor market, acting as a substitute for more routine occupations and, on the other hand, complementing the more skill-intensive occupations. This phenomenon is called job polarization, or routine-biased technological change (RBTC) (Autor and Dorn, 2013, Acemoglu and Autor, 2011). Deming *et al.* (2015) show that in the context of permanent technological advancement, occupations that demand social skills are those that prove more difficult to automate, since they are qualifications based on tacit knowledge, and, therefore, computers are not good substitutes. Their teamwork model argues that workers do not perform equally in different tasks, therefore, teamwork increases productivity via comparative advantages. Deming *et al.* (2015) found positive returns to social skills, and also that there is complementarity between cognitive and social skills. In addition, the literature documents the positive effect of "non-cognitive" skills, including social ones, in the labor market (Borghans *et al.*, 2014, Weinberger 2014).

Finally, the possible non-linearity of the skill wage premium and the urban-size skill wage premium is analyzed through the estimation of models that allow for the existence of distinct returns at different points of the skill distributions. The same *Equation (1)* was estimated with dummies for the quintiles of skill intensity of occupations. Results are presented in Table A1 in the appendix and displayed in figures 3 and 4. Figure 3 presents skill wage premium for each quintile, in relation to the first quintile. The horizontal lines are referred to the right axis, and portray the average premium estimated in Table 3. The columns present the difference in wage levels in relation to the lowest 20% occupations in skill intensity, and all differences are statistically significant. It is clear that wage premium increase as cognitive and social skill intensity increases, and

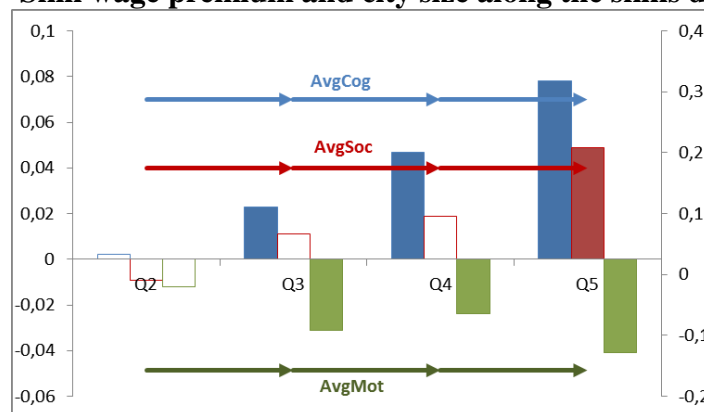
more strongly so for the former. Things are not the same for motor skills, for the negative premium is larger for the upper part of the distribution. Although statistically significant, the difference between the lowest 20% and the next 20% is small. As the occupations become more and more motor skill intensive, the negative wage premium is larger, but the differences between quintiles 3, 4 and 5 are small.

Figure 3 - Skill wage premium along the skills distribution



Source: Organized by the authors.

Figure 4 – Skill wage premium and city size along the skills distribution



Source: Organized by the authors.

The analysis of how the skill wage premium varies with city size, or the urban-size skill wage premium, is presented in Figure 4, in which only the dashed columns indicate statistically significant differences. As in the previous figure, the horizontal lines indicate the coefficient of the interaction between skill and an urban size estimated in Table 3, and are referred to the right axis. On average, the urban-size cognitive wage premium is positive, as presented before, and increases along the distribution of skills. The more intense the cognitive skills, the higher the returns associated with urban population, but this is only statistically significant from the third quintile on. That is, only the top 60% occupations, in terms of cognitive skills, present wage returns to urban size. The same holds for social skills, but only the 20% top social occupations present wage returns to urban size. Again, things are different for motor skills, since, on average, the coefficient associated with the urban-size premium of this skill is negative, and more intensively so as the skill intensity increases. Contrary to the other skills, the

variation is not monotonic, as the value for the fourth quintile is between the values for the third and fifth. When all skills, in quintiles, are estimated simultaneously, the results hold, but the coefficients are typically smaller, and motor skills are only significant for the upper quintile.

5. Panel Results

The previous section analyzed the relationship between the hedonic price of the skills and urban agglomeration, having as reference the OLS model for 2013. In this section, we present evidence on the relationship between skills, wages, and urbanization after controlling for the spatial sorting of workers. Modeling spatial sorting is related to non-observable individual heterogeneity. In this perspective, panel data is needed so that the fixed effects model controls the unobserved time-invariant characteristics of workers. In addition, the possible existence of complementarities between skill mixes and urban size is investigated, as a way to better understand how the different skills profiles relate to urban size. Using annual data for the period 2003-2013, we estimated the following equation:

$$\ln w_{istj} = \phi_{st} \ln(\text{pop}) + \mu_{st} Z_{ij} * \ln(\text{pop}) + \alpha_t X_{ist} + \delta_i Q_{ist} + \varphi_i + \nu_t + \theta_s + \varepsilon_{istj} \quad (2)$$

with w_{istj} standing for the hourly wage of individual i , in occupation j , in LMA s , in year t . X_{ist} is the individual's schooling degree, Q_{ist} is firm characteristics (size and sector), vector Z_t is worker's skills. Fixed effects for individuals, time and macro region⁶ were introduced, and the standard errors were clustered at the level of occupation-year-LMA⁷. The data set is a balanced panel of individuals employed in the private sector in all years. The descriptive analysis of the panel, available in table 4, shows that there are 613,498 individuals, totaling 6,748,478 observations. There is enough variation between firms, occupations, and LMAs to allow for the use of such a rich database.

Table 4 – Descriptive Analysis of Panel Data (2003-2013)

Variation between (%)	
LMAs	16.13
Occupations	64.09
Occupation and LMAs	12.43
Firms	48.46
LMAs, but not Occupation	3.7
Occupation, but not LMAs	51.66
Number of individuals	613,498
Number of observations	6,748,478

Source: Organized by the authors.

⁶ There are five official macro regions in Brazil: North, northeast, center-west, south and southeast.

⁷ To take care of the possible bias in the standard errors of the estimates, as in Bacolod *et al.* (2009) and Groot and DeGroot (2014)

The key results are presented in Table 5. Column 1 does not include skills, and the results indicate the presence of an urban wage premium, even controlling for the spatial sorting of workers, but the intensity of the effect is only a fraction of the OLS estimate for 2013 presented before. This replicates results of other studies that show that sorting is responsible for a large part of the urban wage premium. As for the controls, the results are as expected: wage increases with the size of the firm and with worker's education.

Models 2-4 include interactions between skills and urban size, one skill at a time. The results reinforce the conclusions derived from the cross-section analysis: the effects of cognitive and social skills on wages increase as the size of the city grows. An increase of one standard deviation in the cognitive skill indicator is associated with a 4.8% increase in the urban-size skill wage premium; in the case of social skills, by 3.9%. That is, even after controlling for non-observable time-invariant individual characteristics, there still is an increasing effect of cognitive and social skills as urban size grows. As for motor skills, the results indicate that their effect on wages decreases with the size of the city.

Column 5 presents the results of a model including all variables simultaneously. Urban size and the interaction between size and cognitive and social skills maintain their signs, but with less intensity. Motor skills become positive, although with a very small coefficient. Summing up, the results clearly indicate a positive relationship between cognitive and social skills and urban size, while for motor skills the relationship is either negative or close to zero. In fact, there is no clear evidence that workers with motor skills benefit from the urban wage premium.

Table 5 – Panel Data Results (2003-2013)

Ln (hourly wage)	1	2	3	4	5
Ln(pop)	0.006* (0.001)	0.005* (0.001)	0.005* (0.001)	0.006* (0.001)	0.005* (0.001)
Cognitive*Ln(pop)		0.051* (0.001)			0.038* (0.001)
Social*Ln(pop)			0.041* (0.001)		0.021* (0.001)
Motor*Ln(pop)				-0.019* (0.001)	0.007* (0.001)
High-School	0.031* (0.001)	0.028* (0.001)	0.031* (0.001)	0.030* (0.001)	0.029* (0.001)
College	0.210* (0.003)	0.180* (0.003)	0.193* (0.003)	0.207* (0.003)	0.180* (0.002)
Adjusted R2	0.9338	0.9347	0.9345	0.9339	0.9348
F	42075.07	42387.85	42793.14	40456.81	40085.75
Prob>F	0.00	0.00	0.00	0.00	0.00

Source: Organized by the authors. Standard errors in brackets are clustered at the occupation/year/LMA level. *** p<0.1, ** p<0.05, * p<0.01

As a robustness check, we have estimated *Equation (2)* only with men, following the literature evidence that there are systematic differences between men and women in the

labor market. The results show the same pattern presented before. Another check consisted of substituting total employment for the total population in the regressions, but the results remained, with more intense effects, as expected.

In addition, we sought to investigate whether combinations of skills provide greater payback. The idea is that there may be some combinations of skills that magnify the skill wage premium associated with urban size, since some skill may show complementarity with others. We have inserted interactions of skills with city size in the regressions. The results presented in Table 6 show that the return to the combination of social and cognitive skills increases with city size. No significance was found for the combination *cognitive x motor*, and a negative influence was observed for the *social x motor* combination. Therefore, on top of their isolated importance, cognitive and social skills can have their effect magnified when combined.

Table 6 – Complementarity (2003-2013)

Ln (hourly wage)	1
Ln(pop)	0.005* (0.001)
Cog*Soc*Ln(pop)	0.154* (0.007)
Cog*Mot*Ln(pop)	-0.003 (0.008)
Soc*Mot*Ln(pop)	-0.065* (0.008)
High-School	0.033* (0.001)
College	0.199* (0.003)
ID	Yes
Year	Yes
Sector	Yes
Firm Size	Yes
Macro-Region	Yes
R2 Ajustado	0.9341
F	40799.93
Prob>F	0.00

Source: Organized by the authors. Standard errors in brackets are clustered at the occupation/year/LMA level. *** p<0.1, ** p<0.05, * p<0.01

The urban environment is the locus of opportunities for workers with the right skills, as the results clearly indicate. But why would be some types of occupations more influence than others? Bacolod *et al.* (2010) used contributions from Psychology to study the role of agglomeration and education in the process of skills development, and provided a systemic view of how to understand this mechanism. Skills can be understood as the result of interactions between intelligence and individual traits with the characteristics of environments, notably education and urbanization. Traits can be interpreted as stable characteristics (temperament, personality) that are determined primarily by the genetic factor, and which signal some particular pattern of individual behavior. Intelligence, on the other hand, can be understood from a variety of

perspectives, such as the ability to learn, recognize concepts, or the ability to process information.

Both education and agglomeration influence the manifestation of skills. While education has its relevance widely crystallized, other authors emphasize that formal education is not the only way to promote the development of skills. Informal mechanisms present in agglomerations can also play this role via, for example, the forces of learning. Besides the importance of urbanization in the generation of skills, their model emphasizes that urban agglomeration also impacts wages, since it is directly related to the application of skills in production.

Thus, from the theoretical framework developed by Bacolod *et al.* (2010), it is possible to argue that urban agglomeration is essentially related to cognitive and social skills. In fact, urban forces that enhance the formation of individual abilities, and which, in addition, contribute to the transformation of this skill *pool* into wages, seem to favor more strongly the cognitive and social groups. The explanation for this heterogeneity can be interpreted by looking more closely at what these urban forces are, what their characteristics are, how they operate, and how they interact with different worker profiles.

These forces, which are inherent to the urban environment, can be seen through the micro-foundations of agglomeration economies, that is, *sharing, matching, and learning*. This is the main argument to justify the occurrence of productivity gains resulting from the agglomeration of firms and individuals in space, as well as providing relevant *insights* to understand how distinct this effect is with respect to the profile of the occupations. It is possible to argue that the abler the individual is to absorb the externalities present in large urban centers, the greater his productivity gain. Individuals in occupations with predominantly cognitive and social profiles are the ones to benefit the most from agglomeration.

6. Conclusions

In this paper, we have analyzed how the urban wage premium varies across occupations/workers with different skills. Based on a rich data set of firms and workers, we were able to assign each worker/occupation a score representing the intensity of cognitive, social and motor skills. The geographical analysis of these skills across 369 urban agglomerations, constituting labor market areas, reveals that the relative frequency of individuals with high cognitive and social skills increases as the population grows, whereas, for the motor skills group, it is the opposite.

A cross-section estimation for 2013 allowed identifying a positive association between cognitive and social skills and urban size, and a negative (or non-significant) association between motor skills and population size. This exercise allowed for the determination of a threshold for the existence of returns to social skills, which only appeared for cities close to 800,000 inhabitants. Cognitive skills are present in all city sizes. The panel data analysis reinforced the results, showing that cognitive and social skills provide better wage returns in large cities, while for motor skills the intensity is clearly smaller. These results are robust for different forms of measuring skills and urban size.

An interesting exercise showed that the interaction of cognitive and social skills results in an additional positive effect on the magnitude of the urban-size skill wage premium, but no effect is produced by the interaction of cognitive or social skills with motor

skills. To sum up, exploring a rich dataset, we were able to provide interesting evidence showing that the urban wage premium does not affect equally all occupations/workers, but is quite selective. The advantages of large urban agglomerations are particularly favorable to workers/occupations intensive in cognitive and social skills, although those intensive in motor skills benefit only slightly.

References

- ACEMOGLU, D.; AUTOR D. "Skills, Tasks and Technologies: Implications for Employment and Earnings." *Handbook of Labor Economics*, vol. 4, chapter 12. Elsevier, 1043–1171. (2011)
- ANDERSSON, M., KLAESSON, J.; LARSSON, J. P. "The Sources of the Urban Wage Premium by Worker Skills" (No. 2014/6). Lund University, CIRCLE-Center for Innovation, Research and Competences in the Learning Economy (2014).
- ANDRADE, P. H. P.; GONÇALVES, E.; FREGUGLIA, R. S. "Concentração geográfica de ocupações por regiões metropolitanas brasileiras". *Revista Brasileira de Estudos Regionais e Urbanos*, v. 8, n. 2, p. 155-173. (2015).
- AUTOR, D.; DORN, D. "Inequality and specialization: the growth of low-skill service jobs in the United States". NBER Working Paper Series, n. 15150. (2013).
- AZZONI, C.; SERVO, L.; "Education, cost of living and regional wage inequality in Brazil". *Papers in Regional Science*. V.81. n.2.p. 157-175 (2002).
- BACOLOD, M.P.; BLUM, B. S. "Two Sides of the Same Coin US "Residual" Inequality and the Gender Gap". *Journal of Human Resources*, v. 45, n. 1, p. 197-242, (2010).
- BACOLOD M; BLUM B; STRANGE W., "Skills in the city", *Journal of Urban Economics* 65:136-153, (2009).
- BACOLOD, M.; BLUM, B.S.; STRANGE, W.C. "Elements of skill: traits, intelligences, education, and agglomeration." *Journal of Regional Science* 50.1 (2010): 245-280.
- BARUFI, A. M. B. "Agglomeration economies and labour markets in Brazil". Ph. D. Dissertation. Universidade de São Paulo. (2015).
- BORGHANS, L.; TER WEEL B.; WEINBERG B. A. "People skills and the labor-market outcomes of underrepresented groups." *ILR Review* 67.2: 287-334. (2014).
- CARLSEN, F., RATTSO J., STOKKE H.E. "Urban wage premium and the role of education: Identification of agglomeration effects for Norway." (2012).
- CARLSEN, F.; RATTSO, J.; STOKKE, H.E. "Education, experience, and urban wage premium". *Regional Science and Urban Economics*, v. 60, p. 39-49, (2016).
- COMBES, P.-P., DURANTIN, G., GOBILLON, L., ROUX, S. "Estimating agglomeration economies with history, geology, and worker effects". In: Glaeser, E.L. (Ed.), *Agglomeration Economics*. University of Chicago Press for the NBER, Chicago, IL, pp. 15–65 (2010).
- COMBES, P.P.; DURANTON G.; GOBILLON L. "Spatial wage disparities: Sorting matters!" *Journal of Urban Economics* 63, no. 2 (2008): 723-742.

- COMBES, P.P.; GOBILLON, L. "The empirics of agglomeration economies". In: Handbook of Regional and Urban Economics Vol 5A, North-Holland, Amsterdam, 1st Edition, pp. 247-348, (2015)
- CRUZ, B. O.; NATICCHIONI, P. "Prêmio salarial urbano e a trajetória da desigualdade: Evidências para o Brasil". Texto para Discussão, Instituto de Pesquisa Econômica Aplicada (IPEA), (2012).
- D'COSTA, S.; OVERMAN, H. G. "The urban wage growth premium: Sorting or learning?". *Regional Science and Urban Economics*, v. 48, p. 168-179, (2014).
- DEMING, D. J. "The growing importance of social skills in the labor market". National Bureau of Economic Research. (2015).
- DURANTON, G., PUGA, D., "Micro-foundations of urban agglomeration economies". In: Henderson, J.V., Thisse, J.F. (Eds.), *Handbook of Regional and Urban Economics* 4. Elsevier, Amsterdam, (2004).
- EHRL, P.; MONASTERIO, L. "Historical trades, skills and agglomeration economies". MPRA Paper No. 69829 (2016).
- ECKHOUT, J.; PINHEIRO, R.; SCHMIDHEINY, K. "Spatial sorting: Why New York, Los Angeles and Detroit attract the greatest minds as well as the unskilled". CEPR Discussion Paper No. DP8151. (2010).
- FLORIDA R., MELLANDER C., STOLARICK K., ROSS A. "Cities, skills and wages." *Journal of Economic Geography*, (2011).
- FREGUGLIA, R. S.; MENEZES-FILHO, N. A.; SOUZA, D. B. "Diferenciais salariais inter-regionais, interindustriais e efeitos fixos individuais: uma análise a partir de Minas Gerais". *Estudos Econômicos (São Paulo)*, v. 37, n. 1, p. 129-150, (2007).
- GLAESER, E. L.; D. GOTTLIEB J. D. "The wealth of cities: Agglomeration economies and spatial equilibrium in the United States." No. w14806. National Bureau of Economic Research, (2009).
- GLAESER, E.; MARE, D. "Cities and Skills". *Journal of Labor Economics*, v. 19, n. 2, p. 316-342. (2001).
- GLAESER, E.; RESSEGER, M. G. "The complementarity between cities and skills". *Journal of Regional Science*, v. 50, n. 1, pp. 221-244. (2010).
- GROOT, S.; De GROOT, H. LF. "Estimating the skill bias in agglomeration externalities and social returns to education: Evidence from Dutch matched worker-firm micro-data." (2014).
- GUVENEN, F. "Multidimensional skill mismatch". National Bureau of Economic Research, (2015).
- IBGE. "Arranjos populacionais e concentrações urbanas no Brasil". Coordenação de Geografia. - 2. ed. - Rio de Janeiro: IBGE, (2015).
- INGRAM, B. F.; NEUMANN, G. R. "The returns to skill". *Labour economics*, v. 13, n. 1, p. 35-59, (2006).
- LEVY, F.MURNANE R.J. "The skill content of recent technological change: An empirical exploration". *The Quarterly journal of economics*, v. 118, n. 4, p. 1279-1333, 2003.
- LIU, S. "Agglomeration, Urban Wage Premiums, and College Majors." *Journal of Regional Science* (2016).

- MACIENTE, A.N. "The determinants of agglomeration in Brazil: input-output, labor and knowledge externalities". PhD Thesis, University of Illinois at Urbana-Champaign, (2013).
- MATANO, A.; NATICCHIONI, P. "What drives the urban wage premium? Evidence along the wage distribution". *Journal of Regional Science*, (2015).
- MION, G.; NATICCHIONI P. "The spatial sorting and matching of skills and firms." *Canadian Journal of Economics/Revue Canadienne d'Économique* 42.1p. 28-55, (2009).
- MORETTI, E. "Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data." *Journal of econometrics* 121.1 p. 175-212. (2004).
- OVERMAN, H. G.; PUGA, D. "Labor pooling as a source of agglomeration: An empirical investigation". In: *Agglomeration economics*. University of Chicago Press, p. 133-150 (2010).
- PUGA, D. "The magnitude and causes of agglomeration economies". *Journal of Regional Science*, v. 50, n. 1, pp. 203-219 (2010).
- ROBACK, J. "Wages, rents, and the quality of life". *Journal of political economy*, v. 90, n. 6, p. 1257-1278. (1982).
- ROCHA, R. M.; SILVEIRA NETO, R. M.; GOMES, S. M. F. P. O. "Maiores cidades, maiores habilidades produtivas: ganhos de aglomeração ou atração de habilitados? Uma análise para as cidades brasileiras". In: XVI Encontro Regional de Economia, Fortaleza, julho, (2011).
- ROSENTHAL S.; STRANGE W., "Evidence on the nature and sources of agglomeration economies". In: Henderson, J.V., Thisse, J.F. (Eds.), *Handbook of Regional and Urban Economics* 4. Elsevier, Amsterdam, pp. 2119–2171, (2004).
- SILVA, D. L. G. "Distribuição espacial dos efeitos dos ganhos de aglomeração sobre os retornos à educação no Brasil". 2012. 157 f. Master Thesis – Faculdade de Economia, UFBA, Salvador, (2012)
- WEINBERGER, C. J. "The increasing complementarity between cognitive and social skills", *Review of Economics and Statistics* 96(4), 849–861, (2014).
- WEINSTEIN, A. L. "A Regional Approach to Productive Skills". PhD Thesis. The Ohio State University, (2013).

Adjusted R2	0.5084	0.5109	0.4765	0.4789	0.4580	0.4591	0.5127
F	853.86	1198.19	605.93	676.47	661.51	692.66	932.34
Prob>F	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Source: Organized by the authors. Standard errors are clustered at the occupation/LMA level. *** p<0.1, ** p<0.05, * p<0.01