Department of Economics Working Paper Series
401 Sunset Avenue
Windsor, Ontario, Canada N9B 3P4
Administrator of Working Paper Series: Christian Trudeau
Contact: trudeauc@uwindsor.ca

Network Centrality in Labor Markets and Wage Dynamics

Marcelo Arbex (University of Windsor)
Ricardo Freguglia (Federal University of Juiz de Fora)
Rafael Siano (Federal University of Juiz de Fora)

Working paper 16 - 09

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author or at http://ideas.repec.org/s/wis/wpaper.html.

Network Centrality in Labor Markets and Wage Dynamics*

Marcelo Arbex† Ricardo Freguglia‡ Rafael Siano§

November 23, 2016

Abstract

Network centrality measures how important an individual is for his network overall. We construct a novel measure of network centrality that takes into account a worker's and coworkers' tenure in the same occupation and firm. Using a linked employer-employee data source - the Brazilian Annual Social Information Report (RAIS) - we follow Brazilian workers in the formal labor market and we estimate the effects of a worker's network centrality on their wages. For the period 2008-2013, we identify job networks and calculate a centrality index for workers in the city of São Paulo – Brazil. We present empirical evidence that job network relevance, measure by our centrality index, is important to explain wage differentials in the labor market. The wage returns are shown to be higher for those workers employed in larger companies, as well as for workers who have higher tenure. We also identify a positive relationship between a worker's centrality index and his wage, which suggests that an increase of a worker's relative relevance in his network is associated to higher wages. We believe that our study sheds light on the extent to which a worker's importance in his job network can influence his wage over time.
JEL code: J24, J31, D85
Keywords: Network centrality; Wage differentials; Peer effects; Labor markets.

*We thank FAPEMIG, CAPES and CNPq Brazilian agencies for financial support. We also thank the Brazilian Ministry of Labor for providing us the database.
†University of Windsor, Canada; arbex@uwindsor.ca
‡Federal University of Juiz de Fora, Brazil; ricardo.freguglia@ufjf.edu.br
§Federal University of Juiz de Fora, Brazil; rafaelpsiano@gmail.com

1 Introduction

Networking is a central concept in studies of the labor market. It plays a critical role in job search and in improving the quality of the match between firms and workers. One important feature of networks is the relative centrality of individuals. Centrality is a structural characteristic of individuals in a network. Centrality measures how important an individual is for his network overall. For instance, individuals with high centrality score are often more likely to be leaders, key conduits of information, and be more likely to be early adopters of anything that spreads in a network. In this paper, we construct a novel measure of network centrality that takes into account a worker's and coworkers' tenure in the same occupation and firm. Using a linked employer-employee data source - the Brazilian Annual Social Information Report (RAIS) - we follow Brazilian workers from 2008 to 2013 in the formal labor market and we estimate the effects of a worker's network centrality on their wages. We believe that our study sheds light on the extent to which a worker's importance in his job network can influence his wage over time.

Our paper builds on a growing literature that emphasizes the importance of networks in the labor market. Empirical evidence suggests that a significant fraction of all jobs are found through contacts (see Holzer, 1988; Montgomery, 1991; Topa, 2001; Gregg and Wadsworth, 1996; Addison and Portugal, 2001) and firms use referrals extensively when trying to fill a vacancy. According to the National Longitudinal Survey of Youth (NLSY), more than 85% of workers use informal contacts when searching for a job (Holzer, 1988) and more than 50% of all workers found their job through their social network according to data from the Panel Study of Income Dynamics (Corcoran et al., 1980). Between 37% and 53% of employers use the social networks of their current employees to advertise jobs according to data from the National Organizations Survey (Marsden, 2001) and the Employment Opportunity Pilot Project (EOPP), respectively. According to the EOPP, 36% of firms filled their last opening through a referral (Holzer, 1987). Moreover, the job network literature has shown that social networks have important implications for the dynamics of employment, as well as the duration and persistence of unemployment (Calvó-Armengol and Jackson, 2007).

Our focus on network centrality is in line with evidence presented by Gorter and Ommeren [1999]. The authors state that the networks, as well as the position that the worker is employed in it, are of great importance to their professional success. Furthermore, according to Sparrowe and Liden [1997], the position of the individual within his

network can provide better organizational assimilation to him. Other advantages such as a promotion within a company, for example, may also depend on the position in which the agent occupies in a given network (Burt, 2009). Baldwin et al. [1997], for example, when analyzing the centrality of some agents in one network, found a positive relation between centrality and performance related to the grades of individuals in a network related to an MBA program. For Banerjee et al. [2014], individuals can simply follow the sources of information in a network by identifying other agents that are more central in the network according to the "diffusion of centrality". Testing this intuition by using data from 35 communities in India, they concluded that individuals could indicate precisely those that play a central role in disseminating information in their respective networks. Moreover, these findings indicated agents that are more central in their network than other individuals considered community leaders and geographically central individuals.

Several papers have studied how workers interact with each other and their importance in their labor networks. Notable examples of such papers are Saygin et al. [2014], Beugnot et al. [2013] and Dahl et al. [2014]. Using administrative records of employment for the Austrian private sector, Saygin et al. [2014] noted some networks formed by individuals who worked together in the same company in previous years. To distinguish between information transmission mechanisms, the authors made use of two different perspectives related to the network: from the point of view of the agent looking for a job, and the company's point of view in searching for new workers. For the job applicant, they analyzed how the intrinsic characteristics in his network affect his chances to find new jobs and better wages. For the firm, they analyzed what were the types of workers that could be hired, as the contracting company is linked to another company indirectly through connections formed by individuals who worked together in the past. As a result, they noted that the relationship between former coworkers are crucial for the individual looking for a new job. In addition, 21% of employees who find a new job, do it in a company that is connected to his former workplace. These results highlight the strong role of networks related to the transmission of information among workers. We are the first to specifically trace workers' network centrality over time.

Beugnot et al. [2013]investigates how peer performance and contextual effects can influence individual performance. In this analysis, the authors found that workers are sensitive to peer characteristics, performance, and also to financial incentives. On any network, men are more sensitive than women to financial incentives, and a given increase in financial reward will be translated into gains of productivity. For women, they are

sensitive to peers productivity gains, but to a lesser extent than men. And in a simultaneous setting, women behave as though peer performances were irrelevant, almost as they played in isolation. In other words, male workers seem to thrive on competition although women seem to refuse to play a competitive game. Another study about peer effects was carried out by Dahl et al. [2014], and the authors investigate peer effects in paid paternity leave in Norway using a discontinuous regression. Their findings shows that coworkers and brothers are 11 and 15 percentage points more likely to take paternity leave, respectively, if their peer was induced to take the leave. By the mechanism of information transmission, the individuals were able to know how an employer will react. The estimated peer effect snowballs over time, as the first peer interacts with the second peer, the second peer with the third, and so on.

The contribution of this paper is threefold. First, we develop a centrality index to quantify the relative importance of workers in their networks, specifically taking in to account the number of workers in the job network and their respective tenure on the job. The proposed index intends to fills the gap in the literature for complete networks, since the most common indices that seeks to identify the centrality of an individual are not directly applicable to these types of networks. Workers belong to the same (complete) network if they have the same occupation and work at the same firm in a given time. As every worker is connected to another in a complete network, it is important to define the relative importance of workers in a particular network. One way to differentiate them is by looking at their tenure (time) on the job. Intuitively, ties between workers on the same network get stronger the longer they stay together. For pioneer worker in a network, i.e., a worker that first started at the firm, there are no network benefits since he does not interact with other workers. With additional hiring of other workers for the same firm and occupation, older workers benefit from the arrival of new workers, who in turn benefit by working with more experienced workers. The network centrality index we develop emphasize the fact that a worker's relative importance in his job network is determined not for by his tenure on the job, but in fact by the time he spends with other workers, his occupational coworkers, in the same firm. An underlying assumption here is that tenure is a proxy for work experience, which accumulates over time. For experience to translate into a worker's importance in his job network, job experience has to accumulate during his tenure with other workers on the job. We can show that the relative importance of a worker on his network, measure by his network centrality, will evolve differently over time depending on the interaction of tenure on the job and time

spent with coworkers on the same occupation.

Second, for the period 2008-2013, we identify job networks and calculate a centrality index for each worker in our sample. The analysis is focused on the city of São Paulo - Brazil, and for a longitudinal analysis, only workers who have been tied to any formal job between the years 2008 to 2013 were considered. In São Paulo, we observe more than 100,000 establishments for this period. Occupations of professionals of sciences and arts', 'mid-level technicians', 'workers of administrative services', 'workers in services, trade vendors in stores and markets' and 'workers of production of industrial goods and services' contains over 90% of all workers in this city. Thus, considering this occupational groups we can observe more than 87% of the total number of networks on every year in São Paulo. For these occupations and according to the Brazilian Classification of Occupations (CBO), 479,602 workers in every year were considered in this paper, for a total of 2,877,612 observations.

Finally, we investigate how the relative importance of a worker in his network, i.e., his network centrality, can influence his wage. The mechanism we are considering is that the relative importance of a worker related to their peers in their network, usually considered as a non-observable characteristic in most of the studies in the labor market literature, can influence wage differences. We are in the novel position to identify explicitly the centrality of the workers and, controlling for unmeasured worker-specific characteristics, which are fixed over time, we provide evidence that job network relevance, measure by our centrality index, is important to explain wage differentials in the labor market. The wage returns are shown to be higher for those workers employed in larger companies, as well as for workers who have higher tenure . We also identify a positive relationship between a worker's centrality index and his wage, which suggests that an increase of a worker's relative relevance in his network is associated to higher wages.Furthermore, the addition of controls related to the characteristics of the networks improved the model and those controls were statistically significant at 0.01 level in all the estimations made. To deal with possible problems of endogeneity, we use a fixed effect with instrumental variable approach. As the centrality index and the wage of the worker can be determinated simultaneously, i.e., central workers tend to have a high influence in their network by having a high centrality index, along with a higher wage level. We propose an instrument for the centrality index that is correlated with the index itself, but uncorrelated with the dependent variable. In order to construct a good instrument, we teamed up workers in the same occupation and quintile. In doing so, the centrality index of workers in

different networks with different number of peers can be comparable. Thus, there is now an average representative value for centrality for each quintile related to each occupation. After that, we added up all the mean values of equal quintiles from different occupations, minus the average value from the occupation of interest, and divided that by the number of average values imputed in the sum. According to the estimations made, the improper control for endogeneity presented in the model can underestimate the effect of network centrality of the workers in their wage gains. It may be noted that the coefficient value for the centrality variable obtained by Fixed Effects was smaller than the value obtained by POLS (0.09932 and 0.1238, respectively), and the coefficient value for the centrality index obtained with Fixed Effect with instrumental variable approach (0.269) was higher than the coefficients obtained in both previous methods.It is possible that people who earn more have greater importance among their peers, but regarding the way that we define the centrality index, what matters is not only how many peers a worker have, but how long he works with the same peers. Using the instrumental variable, our results mitigate this individual effect and stress the relative importance of that worker. It is noteworthy that being a worker who get better paid but alone, is less important than a worker who get better paid when having more coworkers.

The structure of the paper is as follows. In the next section we introduce our network definition, discuss the data, and present a detailed descriptive analysis of network centrality. We present the main empirical strategy and results in Section 3 and 4, respectively. And the final section concludes.

2 Network Definition and Data

2.1 Network Definition and Network Centrality

For the Brazilian labor market, we say that workers are connected if they work in the same establishment and have same occupation in a particular year. A network G is a pair (N, g) containing a set of nodes $N = \{1, ..., n\}$ and g matrix nxn with real values, and g_{ij} represents the relation between i and j, where g_{ij}=1 indicates that i and j know each other and g_{ij}=0 otherwise. We assume that g_{ij}=g_{ji}, meaning that the relationship between i and j is reciprocal.

Following Jackson et al. [2008], one can define the centrality measure in various ways: degree centrality (how connected is a node), closeness (how easy a node can reach

another node) and betweenness (how important a node is, in terms of connections with another nodes). According to these usual measures of centrality, in a complete network, all workers of a network would have the same importance, i.e., would all be considered as having virtually the same characteristics, and exert the same influence on each other, in their respective network. However, because of heterogeneity in the characteristics of workers, they can be differentiated from others in their network.

In this study, all networks where workers of the same occupation and of the same firm are inserted are defined as a complete network. All employees connect with each other from the moment in which they relate to that firm and occupation of the Brazilian formal labor market. But how to define the relative importance of the worker in his network? The variable related to the time which the worker is linked to his employment and occupation - that is, their time on the job - can be used to differentiate him among other different workers in a given network. Intuitively, one can infer that the relations between workers get stronger as the time that they are together goes by. A pioneer worker in a network, i.e., the one with most time spent in the firm, in turn, does not receive any gain on the connection with another individual on the network in which it is located if he is the only worker on that network.

As a firm hires more workers for a given occupation, new hires begin to interact with workers already in the firm for some time extending and strengthening their network. Moreover, one can expect that workers with longer tenure have more influence on their network than other workers with less experience in the same occupation and firm. Over time, the relative importance of older workers decreases as connections between the workers in their networks get stronger.

Let N be the set that contains all workers i, ..., j such that $i \neq j$, of any network. Let T represent the time which those workers belonging to N are employed in the same establishment and in the same occupation, thus being in the same network. Formally, the n umber of the connections of a worker i can be defined as:

$$X_i = \sum_{j \in N|i} T_j^i \tag{1}$$

where

$$T_j^i = \begin{cases} T_j & if\ T_j \leq T_i \\ T_i & if\ T_j > T_i \end{cases} \tag{2}$$

With all connections of different workers in any network defined, the analysis can now be focused on creating an index to report the relative importance of each worker in their respective network, taking into account the totality of the individual connections. Formally, the centrality index (CI) of an individual on a complete network can defined as:

$$CI_i = \frac{X_i}{\sum_{j \in N} X_j} \tag{3}$$

The above equation implies that the relative centrality index of an individual depends on their connections (T_i), the connections of their peers ($T_j \in N \setminus i$), and also the size of the network (n), i.e., of how many individuals the network is composed. Thus, we can adapt this statement with the following notation:

$$CI_i = f\,(T_i,\, T_{j \in N|i},\, n) \tag{4}$$

Note that the values for CI are included in the closed interval from zero to one. As we move forward one year, all employees linked to a network will be another year connected to each other, and this fact creates a change in their relative importance on it, i.e., the measured centrality index can be increased or decreased. For the most central workers, as the time passes, their relative importance in the network goes down, and the reverse is also true for workers whose centrality lies among the smallest ones. Consider an increase of a year $\triangle$t = 1. For a given network, we have the number of connections of agent i defined as X_i, and it suffers an increase $\triangle X_i$ = n - 1 over a year. Thus, we can now define that over a year, the new value of X_i, now set for X'_i, can be expressed by:

$$X'_i = X_i + (n-1) \tag{5}$$

Thus, we can say that, over a year, all connections of that network (denominator of the centrality index formula - CI) suffers an increase of the sum of X'_i, ..., X'_j. Formally, all connections of any network is given by:

$$\sum_{j \in N} X'_j = \left[\sum_{j \in N} X_j\right] + [n(n-1)] \tag{6}$$

With the addition of another year, it can be intuitively expected that the centrality index for a worker suffers a decrease for the more central workers and an increase for

workers with smaller centrality in a network.

2.2 Database and Sample

Our empirical analysis is based on the Annual Social Information Report (RAIS) provided by the Ministry of Labour and Employment of the Brazilian government for the years 2008 to 2013. The RAIS database aims the control of labor activity in the country, and also seeks to provide data for the development of labor statistics. By using this database, it is possible a longitudinal monitoring of individuals in the formal labor market in Brazil. The data provide detailed information for over seventy million workers by year in the Brazilian formal labor market.

For this study, the analysis will be focused on the city of São Paulo, a Brazilian city with them most developed financial center and the most populous city in the country, according to the Brazilian Institute of Geography and Statistics - IBGE data, Census of the year 2010. For a longitudinal analysis, workers are identified in the database by their personal identification number (CPF) and only the workers who have been employed in any formal job between the years 2008 to 2013 were considered. In the years considered in this study, the number of establishments in the city of São Paulo does not suffer a major change in its magnitude, and its maximum value is related to 2008, and its minimum can be observed for the year 2010.[1]

Table I presents a detailed description of the variables used in this study. The dependent variable that is used in this study is the logarithm of the worker's wage. We deflate the wages by the Brazilian price index IPCA[2].

[Insert Table I - Variables Description - about here]

The seniority measured in months corresponds to the time that the individual is in the same job. The seniority squared seeks to capture the decreasing marginal returns arising over time. Age is a variable measured in years for each individual, and was included in order to measure the impact of an employee's years of living in their income.

[1] See Siano [2016], in Portuguese, for complete discussion of dataset and centrality index used in this study.

[2] The IPCA is calculated by IBGE for 9 metropolitan areas. The values of the IPCA for the metropolitan region of São Paulo were considered for the preparation of wage deflators in each year. For the year 2013 the deflator is equal 1; for the year 2012, value 0.9425959; for 2011, 0.9001107; for 2010, an amount of 0.8452537; for 2009, 0.7989921; for 2008, the deflator has a value of 0.7642932.

This variable squared seeks to measure the diminishing returns associated with years of workers' life.

The variable that represents how central an individual is when related to his peers in the network is also included in the analysis of this study in order to know if the relative centrality degree of an agent - that is, how much he is important relative to other individuals of the network - exerts some influence on wages observed over the same period. The methodology regarding the calculation of the relative centrality in a complete network, as well as their identification, are presented below. Yet, it is also considered a variable responsible for indicating with how many individuals a worker i is connected, that is, what is the size of his network.

The educational level of the individuals was also considered in this study as a way to try to capture the influence of their level of education in their wages observed in the period. Furthermore, 21 sectors (Section A to Section U) in which the Brazilian labor market is divided according to the National Classification of Economic Activities - CNAE are also considered in the analysis.

The workers interact with each other through connections with peers in their networks, and each agent has significant importance in the networks to which they belong. Such connections are part of the social life and are composed according to mutual interests and characteristics. In this study, it is proposed the identification of these connections among workers by considering the establishment in which they are employed, and the occupation to which they belong and perform their tasks. The identification of establishments in the Brazilian formal labor market can be made by the number of the national register of legal entities (CNPJ) of each firm. As for occupations in Brazil's labor market, these can be distinguished by the Brazilian Classification of Occupations (CBO), established in 2002.

Analyzing the different firms and workers linked in different occupations, one can define that the individuals who work in the same establishment and share the same occupation within that firm, are part of a single network, defined by the combination firm-occupation . Each combination generated in this way is unique, and only defines a single network in which workers who share the same characteristics are inserted.

The option of this study is to focus in the city of São Paulo, capital of the homonymous state, since the same has a developed financial center and lists as the most populous city in Brazil, according to the 2010 Census, conducted by the IBGE. For a longitudinal

analysis possible, as well as control of individual heterogeneity inherent to every individual in the labor market, the unique identification of those workers who have been tied to any formal employment in the entire period of the time included in the analysis of this study was made, and these were the workers considered in this study.

Of all these formal workers in São Paulo, the analysis was focused on workers in occupations in groups 2, 3, 4, 5 and 7, representing the 'professionals of sciences and arts', 'mid-level technicians', 'workers of administrative services', 'workers in services, trade vendors in stores and markets' and the 'workers of production of industrial goods and services', respectively. When considering those occupational groups, over 90% of all workers in the city of São Paulo are included in the analysis.

[Insert Table II - Number of workers by occupational group, São Paulo city - about here]

According to the Brazilian Classification of Occupations established in 2002, the occupational group 2 comprises the occupations whose main activities require for their performance high level of professional knowledge and experience in the physical, biological, social and human sciences. Still, there are professionals of arts and sports in this groups, whose professional practice requires a high level of competence, such as teachers, musicians and others. Its activities consist in expanding the pool of scientific and intellectual knowledge through research; apply concepts and theories for troubleshooting or through education, thus ensuring the systematic dissemination of such knowledge.

Considering now the occupational group 3 of the 'mid-level technicians', it comprises the occupations whose main activities require for their performance expertise and experience in various disciplines of the physical and biological sciences or social sciences and humanities. These activities are to perform technical work related to the application of concepts and methods in relation to the aforementioned spheres.

The occupational group 4 of 'workers of administrative services' comprises two subtypes. The first, includes those who perform bureaucratic work without constant contact with the public, and the second, contains administrative workers on costumer service. The first sub-type includes occupations whose main activities require skills and experience needed to sort, store, compute and retrieve information. These activities consist of performing secretarial work, typing and / or scan and reproduce texts and data in

computers and perform other operations in office equipment. The second sub-type comprises activities with some contact with customers, as performed by library assistants, documentation and post offices, cashiers, attendants, etc.

The occupational group 5 of the 'workers in services, trade vendors in stores and markets', includes occupations whose main tasks require for the knowledge and experience necessary to deliver services to people, protection and security services or selling goods in markets. Such activities consists of travel-related services, housework, restaurants and personal care, protection of people and property and the maintenance of public order, and selling goods in stores and markets.

In addition, the main activities of the occupational group 7 of the 'workers of production of industrial goods and services', require the knowledge to produce industrial goods and services. The group 7 focuses on extractive production workers, construction and industrial production of discrete processes, which mobilize psychomotor and mental skills geared primarily to the form of products.

In this study, considering the database, more than 130,000 networks can be observed in each year, distributed in different occupational groups. For the process of finding the centrality of workers, first, it was verified if the workers were necessarily on the data in all years of analysis. Identified these individuals, one could now identify in which network the workers were in with their peers, and also identify all of their peers for a particular year. Once having made these identifications, the number of connections and the relative centrality index of a worker on a specific year can be calculated. It is noteworthy that the number of peers and the network that the workers are identified may change from year to year, so the measurement of the centrality takes into account only the peers of the workers for the year in which the relative centrality index is being calculated.

For the networks considered in this study, some considerations can be made with regard to the different combinations between establishments and workers linked to them in different occupations. Each combination of firm-occupation shows a single network in which workers with shared characteristics are inserted. However, when it comes to workers related to public administration, we may be underestimating the number of networks in this sector, and overestimating the number of workers employed within a network connected to public administration, defense and social security. Thus, some networks with many workers inside may occur, and this fact can be a source of bias in this study analyzes.

According to the National Classification of Economic Activities (CNAE), the sector designated as 'O' of CNAE 2.0 comprises the activities that are usually carried out by public administration, and these are essentially non-market activities, including general administration (the three levels of government), defense activities, justice, foreign affairs, and the management of compulsory social security system. Thus, when considering workers linked to this sector, hence we are considering a worker i is connected with several other peers with links to public authorities, but not necessarily are relating in the labor market. Thus, a network of entities linked to the government would present a very high number of workers, skewing the analysis. Not considering this workers linked to entities of public administration, defense and social security, more robust results and consistent with reality can be found.

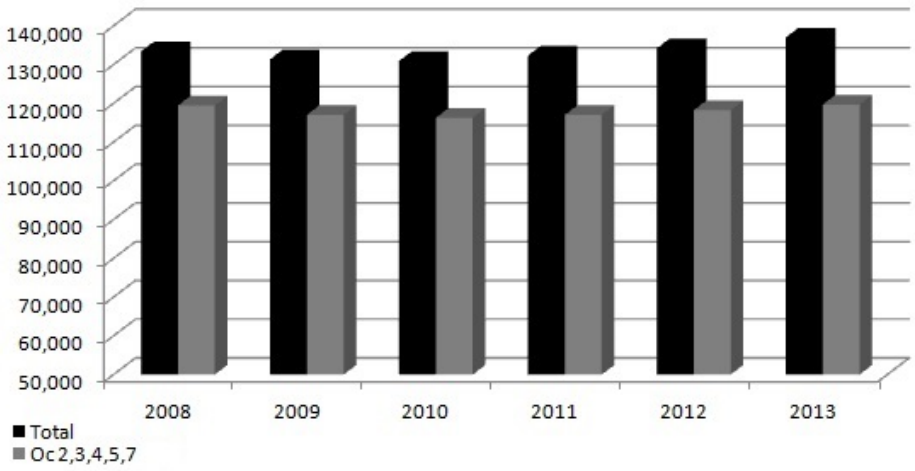
Regarding the number of establishments as well as the networks number for each occupational group considered in this study, from the following table we observe the number of companies, and consequently the maximum number of possible networks in each group considered in this study. The occupations 4 and 5 stand out among the others because they have a larger number of networks, although the occupations 2, 3 and 7 also have a high number of companies in which employees of these occupations are inserted. It is noteworthy that the number of networks on occupations 2, 3, 4, 5 and 7 represents over 87% of all networks in all the years of analysis.

[Insert Table III - Maximum number of networks; Occupational groups - about here]

After the verification of all networks presented in each group, the calculation of the centrality of each worker in his network can be performed for each year. Calculate the centrality of workers for this study will be made with the proposition of a relative centrality index, which will be a function of some individual characteristics related to workers and their calculation method will be explained later on this study.

After this identifying the networks as well as the number of workers in each one of these networks for each year, one can consider only those workers who were exclusively in occupations 2, 3 , 4, 5 and 7, in each year. Some adjustments were also made on the database so that erroneous and missing information were removed from the analysis. Thus, this study now has a total of 479.602 workers in every year, and these workers are necessarily in occupational groups considered in this study from 2008 to 2013. Thus, for workers to be considered in this study, the total number of networks in each years

FIGURE 1 – Total number of networks for the years 2008-2013



can be seen in Figure 1. The total number of networks in the labor market regarding the city of São Paulo is in the range between 130.387 in 2010 to 137.028 in 2013. As noted in the previous chart, the networks related to the occupations highlighted in this study make up much of the total number of networks involved in São Paulo. In 2013, the networks related to such occupations account for 87% of the total, and for the years 2011 and 2012, account for 88% of the total. In the years 2009 and 2010, the networks related to occupations 2, 3, 4, 5 and 7 represent 89% of all networks in the formal labor market of the city of São Paulo, and in 2008, represent 90% of the total.

2.3 Descriptive Statistics

The RAIS database allow us to make a longitudinal follow-up of the workers in Brazilian labor market. As we consider the period between the years 2008-2013 for São Paulo city, the number of workers in each occupational group can be seen in the following table:

[Insert Table IV - Number of workers by occupation for the period 2008-2013 - about here]

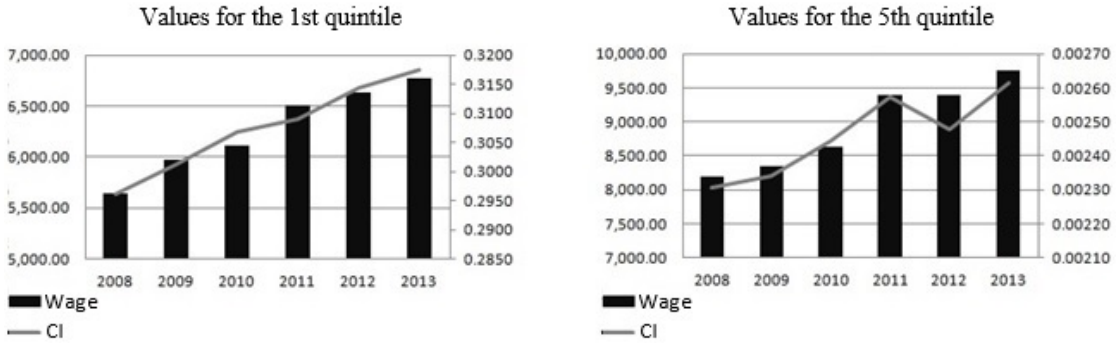
According to the numbers presented about workers in each occupational group for the years 2008-2013, one can observe a constant number for each year of analysis, generating a total of 2,877,612 observations for the period. Thus, a longitudinal analysis is possible in order to control the individual heterogeneity of each worker.

Due to the disparity between the size of the networks considered in this study, a simple comparison of the centrality of workers of different networks do not give us an accurate information about their relationship, as any individual could be the most central in a network with many workers, and have a lower centrality value when compared with another worker on a different network with few individuals. The centrality index of individuals from different networks may be compared to networks having similar characteristics, i.e., networks with few workers can not be compared with other networks with many workers.

Thus, in order to make the wages and the centrality of each worker comparable, the workers were split into five equal parts regarding the number of workers per network, and in the first quintile are 20% of workers that are embedded in networks with few individuals, and in the fifth, there are presented 20% of workers who are part of networks with a high number of individuals. This division allow us to list in every quintile those individuals that are part of different networks, but can be compared due to the similarity in size.

Given the focus of this study, the real wage behavior and the centrality index of each occupation observed from 2008 to 2013 for the city of São Paulo can be observed. The following chart shows the average values of these variables for the occupational group 2, the 'professionals of sciences and arts':

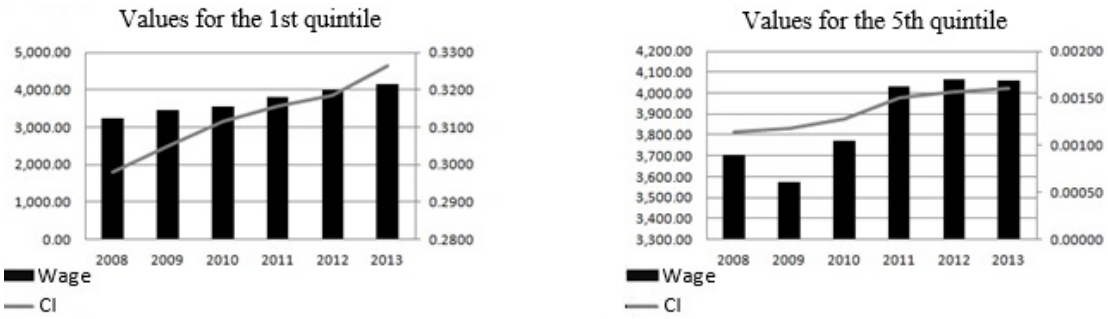
FIGURE 2 – Annual average behavior of real wages and centrality index for group 2



In this occupational group, the mean values of the workers' centrality index exhibit similar behavior to the evolution of the average salary for the first quintile. In the top quintile, there is a wage fall on 2011 to 2012, and this drop is also accompanied by a decrease in the average centrality index of workers discriminated on the 20% of the

population that are inserted in larger networks.

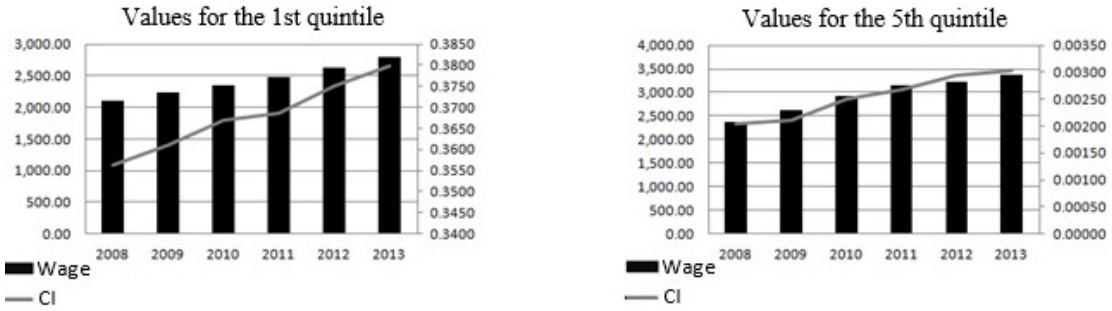
FIGURE 3 – Annual average behavior of real wages and centrality index for group 3



In the occupational group 3 of the 'mid-level technicians', the average behavior of the centrality and the real wage is similar for the 20% of workers inserted on smaller networks. In the top quintile, there is a drop in average salary between 2008 and 2009, and this is not accompanied by a decrease in the average centrality index, but by an increase on it. From 2010 onwards, the average salary shows an increase in its average value, and is also accompanied by an increase in the average centrality of workers in the quintile.

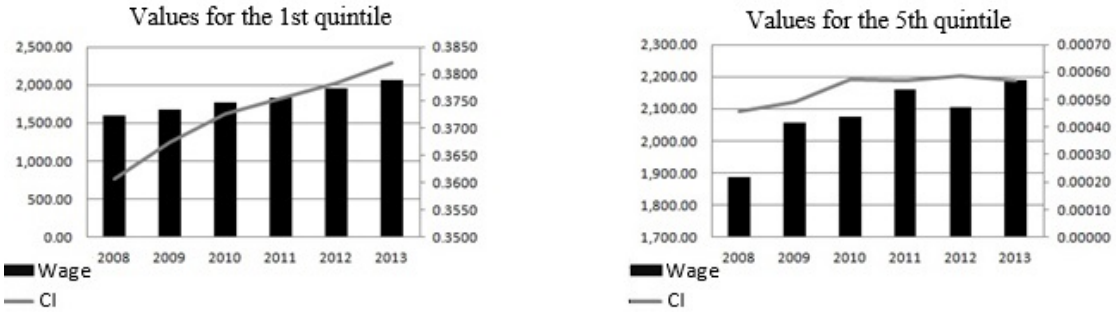
In the following chart, the behavior of the average values of wage and centrality index for the occupational group 4 can be observed:

FIGURE 4 – Annual average behavior of real wages and centrality index for group 4



For the occupational group 4 of the workers of administrative services, there is a similarity on the behavior of the average real wage and workers' average centrality index in both quintiles, and this means that with an increase on the average values for the centrality of a worker are also accompanied by an increase in the average real wages.

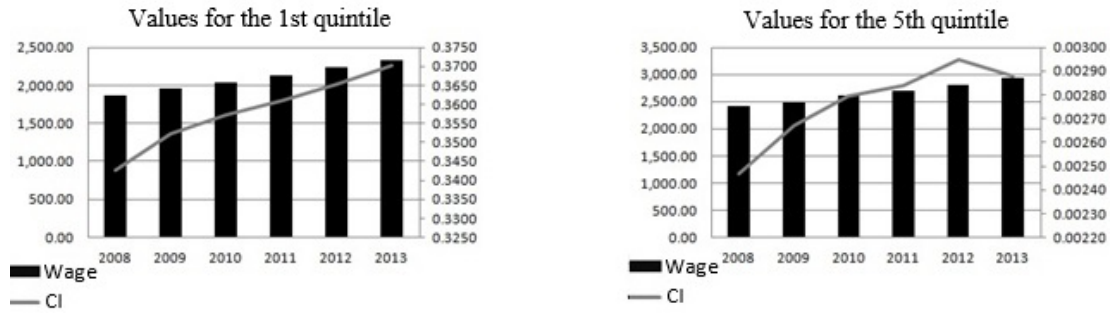
FIGURE 5 – Annual average behavior of real wages and centrality index for group 5



For the occupational group 5 of the workers in services, trade vendors in stores and markets, the average behavior of the variables is similar over the years. As for the top quintile, an increase in the average wage of workers in the years 2008-2010 is also accompanied by an increase in the average centrality of them. From 2010 to 2011, there is a decrease in the average centrality index for workers in this quintile, which is not accompanied by an increase in the average level of real wages. From 2011 to 2012, the opposite occurs, since there is an increase in the average worker centrality and a decrease of the average wage level.

The following chart shows the behavior of the average values of wage and centrality index for the occupational group 7:

FIGURE 6 – Annual average behavior of real wages and centrality index for group 7



For the workers of production of industrial goods and services on group 7, for the first quintile, an increase in the average real wage is followed by an increase in the centrality index over the years 2008 to 2013. In the top quintile, this also takes place between the years 2008 to 2012, and for 2012 to 2013 there is a contraction in the average centrality of workers in the quintile for larger networks, while the average salary level shows an increase over the same period.

Now jointly analyzing all occupational groups, an increase in the average centrality index of the worker, in general, it is also accompanied by an increase in the average level of real wages. It is noteworthy that the centrality index, on average and in absolute terms, are larger for workers in the first quintile, compared with the centrality index of the top quintile of workers. However, this does not mean that a worker with the centrality of 0.3 on the first quintile is 'more important' than a worker with centrality of 0.003 in the top quintile.

A simple comparison between these two numbers does not produce any results, since both workers who possess these centrality values, but are on different networks, can have a central role in their networks. The centrality index proposed in this study is a relative value of the centrality for each worker, and this is a function of their time on job, the time on job of their peers, and the size of their network.

Thus, in order to make a simple descriptive analysis, splitting the workers into groups that differed by the size of their network was extremely important. The centrality of each worker can only be comparable if it is compared with another individual in a different network, but with similar characteristics. Thus, a simple comparison between the index of different workers inserted in the same analysis group can produce credible results, and therefore this results can be open to interpretation.

The descriptive statistics of other variables used in this study are presented in the following table:

[Insert Table V - Descriptive Statistics of the Sample, RAIS, 2008-2013 (1 of 2) - about here]

[Insert Table VI - Descriptive Statistics of the Sample, RAIS, 2008-2013 (2 of 2) - about here]

3 Empirical Strategy

The main goal of the econometric approach presented in our paper is to capture the relationship between wage differentials and the network centrality of a worker in Brazilian formal labor market during the period 2008-2013. We intend to identify the

extent to which a worker's importance in his job network can influence his wage over time. However, several challenges to understand this relationship in Brazil are present. For example, the network centrality that may affect worker wage can be endogenous, since individual unobserved abilities are correlated to the importance of a worker in his job network. In other way, due to heterogeneity of the agents, the estimated results may be influenced by unobserved skills inherent to each individual. This unobservable factor can positively or negatively affect the explanatory variables, overestimating or underestimating the coefficients.

Beginning with a multivariate Mincerian equation (Mincer, 1974), we can obtain the interpretation of the coefficient as the marginal return in terms of income variation. According to the equation bellow, the dependent variable is the logarithm of the individual's salary in its cross-section in a given period of time (Y_{it}), and the explanatory variables can be subsumed in the vector corresponding to the socioeconomic characteristics of individual i at time t (X_{it}: age, age squared, seniority, seniority squared, education, industry occupation, year dummies).

$$Y_{it} = \beta_t X_{it} + e_{it}$$

In this study, the functional form of the model will have as a dependent variable the logarithm of the wage *(lnY)*. The controls of individual features can be described as: age, age squared, seniority, seniority squared, educational level, industry and occupation; such controls, plus the year dummies, are subsumed in the vector *P*. The seniority is measured in months. It corresponds to the workers' time on the same job, and the seniority squared seeks to capture diminishing returns arising over time. Age is a variable measured in years, and was included in order to measure the impact of an employee's years of living in their income. This variable squared seeks to measure the diminishing returns associated with years of one's life. For the worker's industry and occupation, it was followed the National Classification of Economic Activities (CNAE) and the Brazilian Classification of Occupations (CBO), respectively. The control related to the firm's characteristics refers to the size of the establishment, called *F*. Wage differentials related to the importance of the individual in the network, that is, the centrality, can be appointed by *CI*. To the wage differentials linked to the size of the network in which the agent is inserted, it is considered the *size* variable. Formally, the model proposed in this study can be submitted from the following functional form:

$$lnY_{it} = \alpha + \beta P_{it} + \varphi F_{it} + \tau CI_{it} + \lambda worker\,by\,network_{it} + u_{it}$$

In the proposed model, u_{it} is the random error term. In time, if the data used in this study does not define longitudinal information about each individual, estimating the equation above could be made with the Pooled Ordinary Least Squares method (POLS).

However, if we estimate the model proposed earlier by POLS, the results obtained for the estimators may be skewed, since this method do not have the correct treatment for individual heterogeneity. The estimators obtained from this method can be influenced by other unobserved factors inherent to each individual.

With the data from the Annual Social Information Report (RAIS), we are allowed to perform a longitudinal analysis via panel. A panel is characterized when observations about characteristics of the individuals can be made throughout time. With the data organized into a panel, it is possible that unobservable factors (motivation, innate skills, etc.) that influence some characteristics of individuals can be controlled through econometric methods. In this study, methods of Fixed Effects (FE) and Random Effects (RE) will be considered in the analysis, and subsequently will be compared to check what is the method that best fits the proposed model and the data considered. Thus, one can now consider the model in this study, with a new variable "c_i", which does not vary in time and is designated as a way to treat the endogeneity in the regression model:

$$lnY_{it} = \alpha + \beta P_{it} + \varphi F_{it} + \tau CI_{it} + \lambda worker\,by\,network_{it} + c_i + u_{it}$$

In order of having a comparison between methods for the treatment of non-observable effect on a panel, one can use the Hausman test (1978), which has null hypothesis that the estimators of Fixed or Random effects are consistent, but only the estimators of Random Effects are efficient. As an alternative hypothesis, the test states that only the Fixed Effects estimators are consistent and efficient, respecting the level of statistical significance.

In this study, the FE method can be placed in evidence, since the average effect of unobserved factors on wages can be estimated from a longitudinal monitoring of workers, that is, the time path of the agent can be analyzed. If there is some variance in the unobserved factors of individuals over time, c_i and the other independent variables, such as CI and $worker\,by\,network$, for example, are correlated. Therefore, the results

obtained by OLS estimations may be biased. Considering the period of time between the years 2008 to 2013, it is assumed that the unobserved factors related to individuals do not vary over time. With the operation of the difference in individual values with average values, the co-variate c_i can be eliminated, resulting in the acquisition of the Fixed Effects estimator.

With FE, one can treat the bias of omitted variables there are constant over time, but endogeneity can persist on the estimations. Endogeneity can still come by simultaneity, when two variables are co-determinated, with each affecting the other. In our model, the centrality index and the wage of the worker can be determinated simultaneously, i.e., central workers tend to have higher wages, the same way that better payed workers may have a greater influence in their network. In other words, these workers tend to have a great influence in their network by having a great *CI,* along with a higher wage. One way to get consistent and unbiased estimators when having simultaneity is to instrumentalize the endogenous variable *CI*. The instrument for the centrality index must be correlated with the index itself, and uncorrelated with the dependent variable. An instrument must be exogenous and relevant to the model in order to be considered a good instrument for centrality.

In order to construct a good instrument for the centrality index, we teamed up workers in the same occupation and quintile. In doing so, the *CI* of workers in different networks with different number of peers can be comparable. Thus, there is now an average representative value for each quintile related to each occupation. After that, we added up all the mean values of equal quintiles from different occupations, minus the average value from the occupation of interest, and divided that by the number of average values imputed in the sum.

The instrument can be defined as follows, with occupation defined as *i, ..., n* ; and each quintile defined by *j*:

$$CI.IV_{i,j} = Mean\left(\sum \overline{CI}_{n-i,j}\right)$$

The instrumental variable is correlated with the *CI* of the workers because once it is considered the values for each quintile and for each occupation, one can state that for workers in the same occupation and quintile, the measure of centrality instrumentalized by the IV correlates with the centrality of the worker itself. On the same way, the

instrument is not correlated with the dependent variable since that we considered for the occupational group *i,* for example, all the others occupational groups, i.e., *n* - *i.*

Once the instrument is defined, the estimation can be conducted via 2SLS, with the first stage defined as:

$$CI_{it} = \alpha + \beta P_{it} + \varphi F_{it} + \tau CI.IV_{it} + \lambda worker\,by\,network_{it} + c_i + u_{it}$$

Similarly, the second stage can be described as follows:

$$lnY_{it} = \alpha + \beta P_{it} + \varphi F_{it} + \tau \widehat{CI_{it}} + \lambda worker\,by\,network_{it} + c_i + u_{it}$$

4 Results

This section will present the main results of the empirical analysis related to this study. First, will be presented the results of the estimation of the model by the method of Pooled Ordinary Least Squares (POLS). Subsequently, the estimates made by the methods of Fixed Effects (FE) and Random Effects (RE) will also be presented.

In the estimates made by POLS, some specifications for the model were considered. In the first model specification were considered only the controls of individual characteristics. In the second specification, in addition to individual characteristics, were also incorporated specific controls related to firm characteristics. In the third specification, all these controls are also considered, and this specification also has a variable that measures the number of pairs related to a worker i at time t inserted in a particular network, and another variable that measures how important a worker is in relation to their peers, i.e., the centrality index for each worker. Moreover, it is noteworthy that robust methods for the treatment of heteroskedasticity were used.

[Insert Table VII - Results for POLS, period 2008-2013 - about here]

By POLS, it can be seen that the second specification, that has controls for individual characteristics and also controls for firm characteristics, presented a greater degree of adjustment than the first specification, which only has controls for individual characteristics. The Akaike (AIC) and Schwarz (BIC) info criterion also had improved in the

second specification. This result corroborates the fact that the intrinsic characteristics of each establishment are also important in explaining wage differences between workers, and therefore can not be disbelieved as their use for the estimates.

Analyzing the coefficients obtained from the estimation of the third specification by POLS, in which controls related to networks of workers characteristics are also considered, it can be noted that the degree of adjustment increased. Still, the criteria of Akaike (AIC) and Schwarz (BIC) also improved with the addition of these controls for the characteristics of the networks.

As for the coefficients of the third specification, it can be said that wage gains are higher for those workers that are allocated in larger companies, since the coefficients for the variables *dtamestab1* and *dtamestab6* has a negative value, and are significant to the level of 1%. Still, those workers who have more time on the same job also observe higher wage gains, but getting diminishing returns over the years, since the negative and significant coefficient of 0.01 for the variable $senior^2$ (for seniority).

Regarding the result related to the coefficient of the variable *worker_by_net*, it is statistically significant at 0.01, but with a negative sign and low magnitude. The adoption of this variable is important so that the effect of the number of peers of a worker is not incorporated by the variable that measures the centrality of him in his network. Thus, the result of the coefficient obtained for the centrality of the worker only captures the effect of his relative importance in relation to their connections in the network.

In order to not only observe the wage gains over the years, the *CI* variable related to the centrality of the worker in the network can also be highlighted. With its addition to the estimates, one can observe what is the impact on wage of the worker's relative importance in relation to their peers. This variable is significant at 1% level, and has a positive value of magnitude 0.1238. From this result, it can be said that for the city of São Paulo, one worker who has a more central role related to his peers in his network will have greater wage gains. More specifically, if the relative importance of a worker increase by 1 percentage point, his wage will also increase in the ratio of approximately 12% on average.

However, the results obtained from POLS may be skewed, since in this method there is no treatment due to individual heterogeneity of the workers. The estimators obtained from this method can be influenced by other unobserved factors related to each individual.

From the RAIS database, we are allowed to make an analysis via panel. Once the data is arranged in a longitudinal way, it is possible that unobservable factors that somehow influence some characteristics of individuals can be controlled through econometric methods. Thus, the estimates by the methods of Fixed Effects (FE) and Random Effects (RE) were also carried out in order to place the proper control for unobserved factors.

The variables used in the estimations by Random Effects were the same as those previously considered for the method of POLS. For the method of Fixed Effects, those variables that remained invariant over the analysis period was taken out from the specifications after centering around the average values. Furthermore, in order to motivate the use of this method, the overall, between and within variations can be observed on the following table for the variables considered on the regressions for Fixed Effects.

[Insert Table VIII - Overall, between and within variations - about here]

The within variation takes place in a given individual, while the variation between is the variation among individuals. This distinction is important because the estimators differ in the use of these variations. The Fixed Effects estimator can not be identified if there is no within variation. Thus, to check the suitability of the methods of Fixed and Random Effects, the Hausman test was performed. So that the methods are comparable, the specifications used for the method of Fixed Effects were the same for the Random Effects method.

In the following table the results are presented to the estimates for Fixed Effects and Random Effects, considering also three specifications for each method. Then, the result for the Hausman test is also shown:

[Insert Table IX - RE and FE results, period 2008-2013 - about here]

[Insert Table X - Hausman Test - about here]

The Hausman test has as null hypothesis that the estimators of Fixed or Random Effects are consistent, but only the estimators Random Effects are efficient. As an alternative hypothesis, states that only the Fixed Effects estimators are consistent and efficient, respecting the level of statistical significance. According to the results obtained for the value of the χ^2, it can be said that the null hypothesis of the test is rejected,

showing that the Fixed Effects method is preferable against the method of Random Effects.

By controlling for the unobserved effects by estimating the model by Fixed Effects, a better understanding of the real effect of the explanatory variables on the dependent variable is possible. The value for the intra-group R^2 (within) in the three specifications vary in the range between 0.10 and 0.11, as this amount represents the temporal variation in the real wages of workers that is explained by the temporal variation in the explanatory variables. The value of the degree of adjustment for the estimation of Fixed Effects has its smallest value in the first specification, growing in the second and reaching its highest value in the third, with the inclusion of controls related to the individual's peers, as well as their relative importance in the network.

The Akaike's criterion (AIC) and Schwarz (BIC) experienced an improvement from the first to the second specification, when there are - in addition to the controls for the individual characteristics - the controls for the characteristics of the firms. The best values for these criteria can be observed in the third specification, when the controls related to the characteristics of the networks are also considered in the estimates. Therefore, this specification stands out compared to the others.

As for the results obtained with the third specification by Fixed Effects, it can be seen the confirmation of diminishing wage returns in relation to the worker's time on job over the years, since the variable of seniority has a positive sign and a negative sign for its square.For the variables related to the characteristics of the establishment, the coefficients obtained by Fixed Effects are presented in line with the coefficients obtained by POLS. In addition, the variables *dtamestab7* and *dtamestab8* became statistically significant at 0.01 through the estimates for Fixed Effects.

The variable related to the number of peers of a worker in his respective network called *worker_by_net*, it continues to possess statistical significance at a significance level of 0.01. The control of the number of peers of a worker is very important in order to get this effect not embedded in the variable that measures the centrality of the worker. Thus, it is possible that the *CI* variable related to the worker's centrality index only captures in its coefficient the effect of the relative importance of the worker related to their peers in the network.

Despite the observation of wage gains earned by employees over the years, the *CI* variable is also considered in order to observe how the centrality of the worker may be

relevant in determining their wage. When considering this variable as one of the model's controls, it can be seen what is the impact that the worker's relative importance in relation to their peers has on wage gains, not only the wage returns attached to a long time of employment. With the estimation of the third specification by Fixed Effects, such variable presented to be significant at 1%, and has a positive value of 0.09932.

It may be noted that the coefficient value for the centrality variable obtained by Fixed Effects is smaller than the value obtained by POLS (0.09932 and 0.1238, respectively). This difference of the coefficients is related to the unobserved heterogeneity inherent to the workers of the Brazilian labor market. Moreover, one can observe a positive relationship between the centrality of the workers with their wage gains, even after the adoption of the method of Fixed Effects for the treatment of the unobserved characteristics related to the agents.

Although we have treated the bias regarding the unobserved effects by using the FE estimator, the endogeneity persist if the centrality index (CI) and the wage of the worker were determinated simultaneously. This fact implies that central workers tend to have higher wages, the same way that better payed workers may have a greater influence in their network, i.e., these workers tend to have a great influence in their network by having a great *CI,* along with a higher wage. One way to treat the bias from the estimators is to instrumentalize the endogenous variable, *CI.*

The instrument for the centrality index must be correlated with the index itself, and uncorrelated with the dependent variable. An instrument must be exogenous and relevant to the model in order to be considered a good instrument for centrality. The Two-Stage Least Squares (2SLS) estimation with Fixed Effects was carried out in order to make the proper control for the endogeneity caused by the simultaneously determination of CI and the wage level. The instrumental variable $CI.IV$ took place on the estimation by 2SLS, as we show:

[Insert Table XI - Two-stage least squares (2SLS) with Fixed Effects - about here]

According to the first stage of the estimation, we have evidence that the instrumental variable is correlated with the CI of the workers. Since we considered the values for each quintile and for each occupation, for workers in the same occupation and quintile the measure of centrality instrumentalized by the IV correlates with the centrality of the worker itself. The correlation of the instrument and dependent variable is close to zero

(-0.06, precisely), evidencing that the instrumental variable constructed would be a good instrument for network centrality in our model.

Regarding the second stage of the estimation, *CI* was now instrumentalized by all the exogenous variables considered in our model, including the instrumental variable *CI.IV*. It may be noted that the coefficient value for the centrality variable obtained by Fixed Effects was smaller than the value obtained by POLS (0.09932 and 0.1238, respectively), and now the coefficient value for *CI* obtained by 2SLS with Fixed Effects (0.269) is higher than the coefficients obtained in both previous methods. As stated by the estimations made, the improper control for endogeneity presented in the model can underestimate the effect of network centrality of the workers in their wage gains. It is possible that people who earn more have greater importance among their peers, but regarding the way that we define the *CI,* what matters is not only how many peers a worker have, but how long do you work with the same peers. By using the instrument in our model, we mitigate this individual effect and we stress (so increasing the coefficient) the relative importance of that worker. It is noteworthy that being a worker who get better paid but alone, is less important than a worker who get better paid when having more coworkers. After using the instrument, the coefficient of *CI* on our model illustrates exactly what is important is not just how many peers you have or how much working time you have, but how long the employee is working with a group of other workers.

After all adjustments made in order to find the best model for a better and more realistic analysis of the Brazilian labor market, it can be said that, on average, for workers in occupations 2, 3, 4 , 5 and 7 in São Paulo city, and regarding the Fixed Effects method, if the relative importance of a worker compared to his peers increases by 1%, his earnings will also increase about 10%. Thus, when we add an instrument for network centrality of a worker in our model, regarding the estimations made by Two-Stage Least Squares with Fixed Effects, the earnings of the workers will increase about 26% if the relative importance of a worker compared to his coworkers increases 1%. In other words, an increase in the relative importance of workers related with their peers in the network has, on average, a positive impact in their real wage gains.

5 Conclusion

This study identifies the connections between individuals from the Brazilian labor market, and also identified the measure of centrality of some workers in their respective

networks. Measuring this centrality, it was possible that the identification of the importance of a worker related to their peers in their respective network were identified. After this identification, it could then relate this fact to some other individual and firm characteristics, focusing mainly on the analysis of how the relative degree centrality of an agent can influence, along with some other controls, observed wage of the workers.

The relative centrality index proposed in this study fills the gap of the theoretical study for complete networks, since the usual indices that seeks the identification of the centrality of an individual fail when applied to these types of networks. Following the network definition proposed in this study, workers in the same establishment and in the same occupation are defined in a network, and thus, connect from the moment in which they are linked to the company. In order to differentiate them as to their relative importance related to their peers, the time in which the worker is employed was used to determinate such differentiation.

From the identification of the worker centrality , it was observed how the relative importance of a worker related to their peers in their network can influence wage gains. In order to get only the effect of the centrality of the worker on real wages captured, the adoption of controls for individual characteristics and establishment characteristics were considered in the model. The results were obtained through econometric methods aimed at treating individual unobserved heterogeneity, and the method of Fixed Effects could be distinguished from the others as the one that most fits the data and the proposed model. The addition of controls related to the network characteristics improved the model and those controls were statistically significant at 0.01 level of significance in all estimations made. The value of the coefficient for the centrality of the worker had a positive value in all estimations of models by different methods, leading to an interpretation that an increase in the centrality index of a worker would also lead to a high wage gain.

For the treatment of unobserved characteristics related to the heterogeneity present in the Brazilian labor market, the method that presented the best criteria along with the best degree of adjustment was the Fixed Effects. In this model, the wage returns are shown to be higher for those workers employed in larger companies, as well as for workers who have higher employment time. Furthermore, the centrality of the worker also performed positively related to wage gains by the individuals. Thus, if the relative importance of a worker increases, their wage will also increase. Although we have treated the bias regarding the unobserved effects by using the FE estimator, the endogeneity persisted, caused by the simultaneously determination of the centrality index (CI) and

the wage of the workers. Regarding that, we treated the bias from the estimators by using an instrument for network centrality. The results from 2SLS with Fixed Effects showed that the improper control for endogeneity presented in the model underestimated the effect of network centrality of the workers in their wage gains.

In time, the relevance of this study hangs next to the proposition of the relative centrality index for complete networks and its application in an econometric model. This application allowed us to identify the effect of the importance that each employee has in the network related to their real wage. As for a future extension, the sample cut may be revised upwards, in order to consider a larger number of employees in the Brazilian labor market in the analysis of centrality.

Attachments

Variables	Description
Dependent variable	
Logarithm of the real wage	Logarithm of the real wage observed in December (deflated by the IPCA)
Explanatory variables	
Controls for individual characteristics	
Age	Individual's age
Age^2	Individual's age squared
Seniority	Length of the time in the current job in months
Seniority ^{2}	Length of the time in the current job in months squared
Gender	Dummy (1,0) of gender - 1 for male, 0 for female
Centrality Index (CI)	Centrality Index (CI) of an indivudual in a network
Workers by network	Total number of individuals in the same network
Educational level	
Illiterate	Dummy (1,0) of educational level
5th year incomplete	Dummy (1,0) of educational level
5th year complete	Dummy (1,0) of educational level
6th to 9th year incomplete	Dummy (1,0) of educational level
Complete primary	Dummy (1,0) of educational level
Incomplete high school	Dummy (1,0) of educational level
Complete high school	Dummy (1,0) of educational level
Incomplete college	Dummy (1,0) of educational level
Complete college	Dummy (1,0) of educational level
Master's degree	Dummy (1,0) of educational level
Phd	Dummy (1,0) of educational level
Sector	
Section A	Dummy (1,0) - 1 for Agriculture, Livestock, Forestry Production, Fisheries and Aquaculture
Section B	Dummy (1,0) - 1 for Extractive Industries
(continues)	
Section C	Dummy (1,0) - 1 for Manufacturing industry

Section D	Dummy (1,0) - 1 for Electricity and Gas
Section E	Dummy (1,0) - 1 for Water, Sewage, Waste Management activities and decontamination
Section F	Dummy (1,0) - 1 for Construction
Section G	Dummy (1,0) - 1 for Trade; Vehicle repair
Section H	Dummy (1,0) - 1 for Transportation, Storage and Mail
Section I	Dummy (1,0) - 1 for Accommodation and Food
Section J	Dummy (1,0) - 1 for Information and Communication
Section K	Dummy (1,0) - 1 for Financial Services, Insurance and Related Services
Section L	Dummy (1,0) - 1 for Real Estate Activities
Section M	Dummy (1,0) - 1 for Professional, scientific amd technical activities
Section N	Dummy (1,0) - 1 for Administrative Activities and Complementary Services
Section O	Dummy (1,0) - 1 for Public administration, defense and social security
Section P	Dummy (1,0) - 1 for Education
Section Q	Dummy (1,0) - 1 for Human Health and Social Services
Section R	Dummy (1,0) - 1 for Arts, Culture, Sport and Recreation
Section S	Dummy (1,0) - 1 for Other Activities and Services
Section T	Dummy (1,0) - 1 for Home Services
Section U	Dummy (1,0) - 1 for International Organizations and other extraterritorial instituitions
Occupation	
Occupational Group 1	Dummy (1,0) - 1 for Senior members of government, leaders of public organizations
Occupational Group 2	Dummy (1,0) - 1 for Professionals of Sciences and Arts
Occupational Group 3	Dummy (1,0) - 1 for Mid-level technicians
(continues)	
Occupational Group 4	Dummy (1,0) - 1 for Workers of administrative services
Occupational Group 5	Dummy (1,0) - 1 for Workers in services, trade vendors in stores and markets

Occupational Group 6	Dummy (1,0) - 1 for Agricultural workers, forestry and fishery
Occupational Group 7	Dummy (1,0) - 1 for Workers of production of industrial goods and services
Occupational Group 8	Dummy (1,0) - 1 for Workers of production of industrial goods and services
Occupational Group 9	Dummy (1,0) - 1 for Workers in repair and maintenance services
Controls for establishment characteristics	
Size of the firm	
0 employee	Dummy (1,0) for the size of the firm
Up to 4 employees	Dummy (1,0) for the size of the firm
5 a 9 employees	Dummy (1,0) for the size of the firm
10 to 19 employees	Dummy (1,0) for the size of the firm
20 to 49 employees	Dummy (1,0) for the size of the firm
50 to 99 employees	Dummy (1,0) for the size of the firm
100 to 249 employees	Dummy (1,0) for the size of the firm
250 to 499 employees	Dummy (1,0) for the size of the firm
500 to 999 employees	Dummy (1,0) for the size of the firm
1000 or more employees	Dummy (1,0) for the size of the firm
Year	
_2008	Dummy (1,0) of year
_2009	Dummy (1,0) of year
_2010	Dummy (1,0) of year
_2011	Dummy (1,0) of year
_2012	Dummy (1,0) of year
_2013	Dummy (1,0) of year

TABLE II – Number of workers by occupational group, São Paulo city

Group	Year							
	2008	2009	2010	2011	2012	2013	Total Obs.	%
1 - Senior members of government, leaders of public organizations	60,220	62,496	66,344	70,833	77,624	86,026	423,543	5.47
2 - Professionals of Sciences and Arts	178,011	181,383	187,699	192,135	196,982	194,091	1,130,301	14.59
3 - Mid-level technicians	332,846	331,629	325,339	324,432	322,719	318,801	1,955,766	25.24
4 - Workers of administrative services	262,781	256,760	256,256	253,699	247,992	249,375	1,526,863	19.7
5 - Workers in services, trade vendors in stores and markets	283,425	287,846	283,953	280,337	277,396	275,170	1,688,127	21.78
6 - Agricultural workers, forestry and fishery	2,576	2,581	2,636	2,674	2,721	2,701	15,889	0.21
7 - Workers of production of industrial goods and services	131,774	129,140	129,106	127,301	126,197	125,857	769,375	9.93
8 - Workers of production of industrial goods and services	22,014	22,341	22,749	22,780	22,461	22,325	134,670	1.73
9 - Workers in repair and maintenance services	17,999	17,379	17,528	17,300	17,236	17,241	104,683	1.35

TABLE III – Maximum number of networks; Occupational groups

Group	Year					
	2008	2009	2010	2011	2012	2013
1 - Senior members of government, leaders of public organizations	6,683	6,971	7,206	7,776	8,367	9,640
2 - Professionals of Sciences and Arts	14,578	14,498	14,927	15,771	16,745	16,848
3 - Mid-level technicians	14,390	14,127	14,025	14,241	14,412	14,814
4 - Workers of administrative services	29,610	29,016	28,673	28,499	28,556	28,849
5 - Workers in services, trade vendors in stores and markets	38,696	37,539	36,971	36,893	36,900	37,316
6 - Agricultural workers, forestry and fishery	370	364	366	371	391	396
7 - Workers of production of industrial goods and services	22,184	21,870	21,693	21,615	21,632	21,871
8 - Workers of production of industrial goods and services	4,047	4,129	4,189	4,260	4,291	4,331
9 - Workers in repair and maintenance services	2,858	2,873	2,887	2,910	2,927	2,963

TABLE IV – Number of workers by occupation for the period 2008-2013

Group	Year						
	2008	2009	2010	2011	2012	2013	Total Obs.
2 - Professionals of Sciences and Arts	66,530	68,575	71,249	74,682	77,834	79,838	438,708
3 - Mid-level technicians	73,558	74,445	74,155	74,191	73,907	73,921	444,177
4 - Workers of administrative services	114,171	113,397	111,751	110,259	108,826	107,706	666,110
5 - Workers in services, trade vendors in stores and markets	144,869	143,200	141,653	140,406	139,033	137,954	847,115
7 - Workers of production of industrial goods and services	80,474	79,985	80,794	80,064	80,002	80,183	481,502
Total of workers	479,602	479,602	479,602	479,602	479,602	479,602	2,877,612

Variable	Continuous / Categorical	Category	Occupational Group 2		Occupational Group 3	
			Mean / %	Std. Dev.	Mean / %	Std. Dev.
Real wage	Continuous		7,518.53	6,038.77	4,245.85	3,885.53
Age (years)	Continuous		39.48	9.69	38.73	9.53
Seniority (months)	Continuous		107.87	94.45	109.5	92.74
Gender	Categorical	Male	46.44		49	
		Female	53.56		51	
Workers by network	Continuous		169.95	316.56	295.72	536.25
Educational level	Categorical	Illiterate	0		0.005	
		5th year incomplete	0.04		0.36	
		5th year complete	0.05		0.89	
		6th to 9th year incomplete	0.13		1.6	
		Complete primary	0.38		4.33	
		Incomplete high school	0.38		3.66	
		Complete high school	5.11		52.65	
		Incomplete college	4.61		9.7	
		Complete college	83.48		26.6	
		Master's degree	1.62		0.08	
		Phd	4.2		0.13	
Size of the firm	Categorical	0 employee	0.05		0.02	
		Up to 4 employees	0.24		0.66	
		5 a 9 employees	0.89		1.66	
		10 to 19 employees	2.85		3.79	
		20 to 49 employees	7.39		8.93	
		50 to 99 employees	7.55		8.33	

(continues)

100 to 249 employees	13.35	11.68
250 to 499 employees	15.91	12.12
500 to 999 employees	15.11	12.4
1000 or more employees	36.66	40.41

Variable	Continuous / Categorical	Category	Occupational Group 4		Occupational Group 5		Occupational Group 7	
			Mean / %	Std. Dev.	Mean / %	Std. Dev.	Mean / %	Std. Dev.
Real wage	Continuous		2934.799	2,781.73	1,964.12	2,011.49	2,307.67	1,753.89
Age (years)	Continuous		36.74	9.94	39.79	9.63	40.71	10.02
Seniority (months)	Continuous		94.19	82.85	83.78	65.38	92.24	72.84
Gender	Categorical	Male	44.63		66.14		82.56	
		Female	55.37		33.86		17.44	
Workers by network	Continuous		219.08	543.41	752.5	1484.05	173.39	402.72
Educational level	Categorical	Illiterate	0		0.45		0.42	
		5th year incomplete	0.59		5.5		6.28	
		5th year complete	0.59		9.31		10.22	
		6th to 9th year incomplete	2.63		14.01		15.33	
		Complete primary	5.41		20.16		21.97	
		Incomplete high school	4.78		7.85		8.77	
		Complete high school	53.19		36.45		34.94	
		Incomplete college	9.16		1.73		0.88	
		Complete college	22.98		4.52		1.19	
		Master's degree	0.05		0.005		0.003	
		Phd	0.02		0.004		0.002	
Size of the firm	Categorical	0 employee	0.03		0.05		0.03	
		Up to 4 employees	1.67		2.94		1.68	
		5 a 9 employees	3.4		12.97		4.72	
		10 to 19 employees	7.19		9.45		9.14	
		20 to 49 employees	13.43		12.08		15.18	
		50 to 99 employees	11.44		7.57		12.21	*(continues)*
		100 to 249 employees	14.63		7.96		15.33	

250 to 499 employees	12.86	6.25	12.88
500 to 999 employees	9.67	6.38	10.31
1000 or more employees	25.68	34.35	18.52

TABLE VII – Results for POLS, period 2008-2013

Dependent Variable - Logarithm of the real wages						
Variables	(1)	(2)	(3)	(1.r)	(2.r)	(3.r)
Gender	0.2608438***	0.2549554***	0.2528941***	0.2608438***	0.2549554***	0.2528941***
	(0.0006942)	(0.0006861)	(0.0006859)	(0.0007071)	(0.0007027)	(0.0007022)
Seniority	0.0024307***	0.0025364***	0.0024693***	0.0024307***	0.0025364***	0.0024693***
	(0.0000121)	(0.0000120)	(0.0000120)	(0.0000126)	(0.0000125)	(0.0000126)
Seniority^2	-0.0000010***	-0.0000014***	-0.0000013***	-0.0000010***	-0.0000014***	-0.0000013***
	(0.0000000)	(0.0000000)	(0.0000000)	(0.0000000)	(0.0000000)	(0.0000000)
0 employee		-0.1223738***	-0.1688908***		-0.1223738***	-0.1688908***
		(0.0161162)	(0.0161095)		(0.0190493)	(0.0190873)
Up to 4 empl.		-0.3251413***	-0.4079142***		-0.3251413***	-0.4079142***
		(0.0024782)	(0.0027695)		(0.0022257)	(0.0025645)
5 a 9 empl.		-0.1407595***	-0.2159895***		-0.1407595***	-0.2159895***
		(0.0014812)	(0.0018032)		(0.0013712)	(0.0017446)
10 to 19 empl.		-0.2010387***	-0.2607689***		-0.2010387***	-0.2607689***
		(0.0013643)	(0.0015799)		(0.0013405)	(0.0015887)
20 to 49 empl.		-0.1783713***	-0.2280931***		-0.1783713***	-0.2280931***
		(0.0011373)	(0.0013028)		(0.0011403)	(0.0013277)
50 to 99 empl.		-0.1096716***	-0.1522834***		-0.1096716***	-0.1522834***
		(0.0012164)	(0.0013277)		(0.0012406)	(0.0013639)
100 to 249 empl.		-0.0137459***	-0.0506875***		-0.0137459***	-0.0506875***
		(0.0011120)	(0.0012046)		(0.0011477)	(0.0012506)
250 to 499 empl.		0.0297839***	-0.0015216		0.0297839***	-0.0015216
		(0.0011349)	(0.0012056)		(0.0011560)	(0.0012352)
500 to 999 empl.		0.0552513***	0.0293592***		0.0552513***	0.0293592***
		(0.0011516)	(0.0012063)		(0.0011854)	(0.0012464)
1000 or more empl.		-	-		-	-

(continues)

Workers by net.			-0.0000275***			-0.0000275***
			(0.0000004)			(0.0000004)
Centrality Index			0.1238066***			0.1238066***
			(0.0029083)			(0.0029071)
Observations	2,877,612	2,877,612	2,877,612	2,877,612	2,877,612	2,877,612
R^2	0.555	0.566	0.567	0.555	0.566	0.567
R^2 adjusted	0.5551	0.5665	0.5674	-	-	-
AIC	4,478,123	4,403,350	4,397,008	4,478,123	4,403,350	4,397,008
BIC	4,478,690	4,404,032	4,397,716	4,478,690	4,404,032	4,397,716

Note: $^{*}p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. (1) controls for individual characteristics; (2) individual characteristics and firm characteristics; (3) controls for individual and firm carachteristics; controls for number of peers and centrality index. On specifications (X.r), the standard error value is robust.

Variables	Overall		Between		Within		Within variation
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	
Seniority	95.25341	80.74518	95.25341	76.76743	95.25341	25.03108	0.09610
Seniority^2	15,592.99	26,968.26	15,592.99	26,142.04	15,592.99	6,624.34	0.06034
dtamestab0 - 0 employee	0.00037	0.01925	0.00037	0.00793	0.00037	0.01755	0.83052
dtamestab1 - Up to 4 employees	0.01673	0.12827	0.01673	0.11201	0.01673	0.06252	0.23754
dtamestab2 - 5 a 9 employees	0.05787	0.23350	0.05787	0.20943	0.05787	0.10325	0.19552
dtamestab3 - 10 to 19 employees	0.06998	0.25511	0.06998	0.21875	0.06998	0.13126	0.26472
dtamestab4 - 20 to 49 employees	0.11708	0.32151	0.11708	0.28244	0.11708	0.15361	0.22827
dtamestab5 - 50 to 99 employees	0.09356	0.29122	0.09356	0.24864	0.09356	0.15161	0.27103
dtamestab6 - 100 to 249 employees	0.12134	0.32652	0.12134	0.28538	0.12134	0.15866	0.23611
dtamestab7 - 250 to 499 employees	0.11268	0.31620	0.11268	0.27535	0.11268	0.15544	0.24167
dtamestab8 - 500 to 999 employees	0.10057	0.30076	0.10057	0.25543	0.10057	0.15877	0.27869
dtamestab9 - 1000 or more employees	0.30982	0.46242	0.30982	0.43893	0.30982	0.14551	0.09902
Workers by net.	372.81	930.20	372.81	880.53	372.81	299.90	0.10394
Centrality Index	0.11343	0.14742	0.11343	0.13680	0.11343	0.05496	0.13899

Note: The within variation is obtained with the within variance divided by the overall variance

Dependent Variable - Logarithm of the real wages						
Variables	Random Effects			Fixed Effects		
	(1)	(2)	(3)	(1)	(2)	(3)
Gender	0.2381628***	0.2374439***	0.2367292***			
	(0.0015332)	(0.0015099)	(0.0015081)			
Seniority	0.0013373***	0.0013555***	0.0013342***	0.0038110***	0.0037799***	0.0037294***
	(0.0000111)	(0.0000111)	(0.0000111)	(0.0000103)	(0.0000103)	(0.0000104)
Seniority^2	-0.0000004***	-0.0000005***	-0.0000004***	-0.0000049***	-0.0000049***	-0.0000048***
	(0.0000000)	(0.0000000)	(0.0000000)	(0.0000000)	(0.0000000)	(0.0000000)
0 employee		-0.1026159***	-0.1232845***		-0.1411857***	-0.1621430***
		(0.0080829)	(0.0080945)		(0.0083665)	(0.0083743)
Up to 4 empl.		-0.2318657***	-0.2653775***		-0.1931906***	-0.2329984***
		(0.0025020)	(0.0026628)		(0.0029937)	(0.0031148)
5 a 9 empl.		-0.2012975***	-0.2317317***		-0.1694780***	-0.2027748***
		(0.0017642)	(0.0019204)		(0.0022509)	(0.0023587)
10 to 19 empl.		-0.1822884***	-0.2089713***		-0.1490452***	-0.1764368***
		(0.0015046)	(0.0016258)		(0.0019004)	(0.0019856)
20 to 49 empl.		-0.1495793***	-0.1729231***		-0.1179893***	-0.1399887***
		(0.0013176)	(0.0014099)		(0.0016595)	(0.0017254)
50 to 99 empl.		-0.1034482***	-0.1240674***		-0.0868258***	-0.1048661***
		(0.0012943)	(0.0013627)		(0.0015750)	(0.0016261)
100 to 249 empl.		-0.0460240***	-0.0639314***		-0.0475001***	-0.0618469***
		(0.0011907)	(0.0012452)		(0.0014186)	(0.0014607)
250 to 499 empl.		-0.0079471***	-0.0225553***		-0.0225521***	-0.0332811***
		(0.0011310)	(0.0011699)		(0.0012993)	(0.0013300)
500 to 999 empl.		0.0086456***	-0.0013594		-0.0034131***	-0.0101748***
		(0.0009822)	(0.0010043)		(0.0010812)	(0.0010987)
						(continues)
1000 or more empl.		-	-		-	-

Workers by net.			-0.0000196***			-0.0000146***
			(0.0000004)			(0.0000005)
Centrality Index			0.0448126***			0.0993289***
			(0.0025130)			(0.0027389)
Observations	2,877,612	2,877,612	2,877,612	2,877,612	2,877,612	2,877,612
R^2 within	0.1749	0.1769	0.1773	0.1038	0.1068	0.1077
R^2 between	0.5752	0.5875	0.5888	0.2708	0.3101	0.3107
R^2 overall	0.5337	0.5457	0.5469	0.2506	0.2863	0.2868
AIC				-527,881	-537,790	-540,476
BIC				-527,546	-537,339	-540,000

Note: $^{*}p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard error in parentheses;
(1) controls for individual characteristics; (2) individual characteristics and firm characteristics;
(3) controls for individual and firm; characteristics, controls for number of peers and centrality index.

Dependent Variable - Logarithm of the real wages						
Variables	(2)			(3)		
	FE	RE	Dif.	FE	RE	Dif.
Seniority	0.0037799	0.0035673	0.0002127	0.0037294	0.0035265	0.0002029
Seniority^2	-0.00000486	-0.00000370	-0.00000116	-0.00000477	-0.00000363	-0.00000114
dtamestab0 - 0 employee	-0.1411857	-0.1938201	0.0526344	-0.162143	-0.2188054	0.0566623
dtamestab1 - Up to 4 employees	-0.1931906	-0.3113406	0.11815	-0.2329984	-0.3544784	0.12148
dtamestab2 - 5 a 9 employees	-0.169478	-0.2684518	0.0989738	-0.2027748	-0.3066892	0.1039144
dtamestab3 - 10 to 19 employees	-0.1490452	-0.2376532	0.088608	-0.1764368	-0.2703686	0.0939318
dtamestab4 - 20 to 49 employees	-0.1179893	-0.1916283	0.073639	-0.1399887	-0.2194582	0.0794695
dtamestab5 - 50 to 99 employees	-0.0868258	-0.1331866	0.0463608	-0.1048661	-0.1572215	0.0523554
dtamestab6 - 100 to 249 employees	-0.0475001	-0.0613031	0.013803	-0.0618469	-0.081675	0.0198281
dtamestab7 - 250 to 499 employees	-0.0225521	-0.0139744	-0.0085778	-0.0332811	-0.0302029	-0.0030782
dtamestab8 - 500 to 999 employees	-0.0034131	0.0083816	-0.0117948	-0.0101748	-0.0025015	-0.0076733
Workers by net.				-0.0000146	-0.0000209	0.00000631
Centrality Index				0.0993289	0.0718664	0.0274625
Observations		2,877,612			2,877,612	
χ^2		199782.15			200063.64	
Prob $> \chi^2$		0.00			0.00	

	First Stage	Second Stage
	Dependent Variable - Centrality Index	Dependent Variable - Logarithm of the real wages
Variables	(3)	(3)
Seniority	0.0003832***	0.0036468***
	(0.0000001)	(0.0000105)
Seniority²	-0.0000001***	-0.0000046***
	(0.0000000)	(0.0000000)
0 employee	0.0394712***	-0.1820078***
	(0.0014961)	(0.0083893)
Up to 4 empl.	0.1116019***	-0.2810129***
	(0.0005541)	(0.0032447)
5 a 9 empl.	0.0697632***	-0.2395723***
	(0.0004225)	(0.0024594)
10 to 19 empl.	0.0449867***	-0.2032558***
	(0.0003564)	(0.0020499)
20 to 49 empl.	0.0280138***	-0.1580491***
	(0.0003095)	(0.0017597)
50 to 99 empl.	0.0174857***	-0.1171691***
	(0.0002913)	(0.0016437)
100 to 249 empl.	0.0089006***	-0.0691657***
	(0.0002614)	(0.0014683)
250 to 499 empl.	0.0039866***	-0.0370154***
	(0.0002377)	(0.0013329)
500 to 999 empl.	0.0010391***	-0.0114701***
	(0.0001963)	(0.0010999)
1000 or more empl.	-	-
	-	-
		(continues)
Workers by net.	-0.000000546***	-0.0000138***

	(0,00000001)	(0.0000005)
Centrality	0.7099811***	-
Index_IV		
	(0.0005324)	-
Centrality Index	-	0.2690160***
	-	(0.0042005)
Observations	2,877,612	2,877,612
R^2 within	0.4873	0.1063
R^2 between	0.8884	0.3055
R^2 overall	0.8294	0.2816

Note: $^{*}p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard error in parentheses;
(3) controls for individual and firm characteristics, controls for number of peers and centrality index.

References

John T Addison and Pedro Portugal. Job search methods and outcomes. 2001.

Timothy T Baldwin, Michael D Bedell, and Jonathan L Johnson. The social fabric of a team-based mba program: Network effects on student satisfaction and performance. *Academy of management journal*, 40(6):1369–1397, 1997.

Abhijit Banerjee, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson. Gossip: Identifying central individuals in a social network. Technical report, National Bureau of Economic Research, 2014.

Julie Beugnot, Bernard Fortin, Guy Lacroix, and Marie Claire Villeval. Social networks and peer effects at work. *CIRANO-Scientific Publications 2013s-27*, 2013.

Ronald S Burt. *Structural holes: The social structure of competition*. Harvard university press, 2009.

Antoni Calvó-Armengol and Matthew O Jackson. Networks in labor markets: Wage and employment dynamics and inequality. *Journal of economic theory*, 132(1):27–46, 2007.

Mary Corcoran, Linda Datcher, and Greg J Duncan. Most workers find jobs through word of mouth. *Monthly Labor Review*, pages 33–35, 1980.

Gordon B Dahl, Katrine V Løken, and Magne Mogstad. Peer effects in program participation. *American Economic Review*, 104(7):2049–2074, 2014.

Cees Gorter and Jos van Ommeren. Sequencing, timing and filling rates of recruitment channels. *Applied Economics*, 31(10):1149–1160, 1999.

Paul Gregg and Jonathan Wadsworth. How effective are state employment agencies? jobcentre use and job matching in britain. *Oxford Bulletin of Economics and Statistics*, 58(3):443–467, 1996.

Harry J Holzer. Hiring procedures in the firm: their economic determinants and outcomes. Technical report, National Bureau of Economic Research, 1987.

Harry J Holzer. The determinants of employee productivity and earnings: some new evidence. Technical report, National Bureau of Economic Research, 1988.

Matthew O Jackson et al. *Social and economic networks*, volume 3. Princeton University Press Princeton, 2008.

Peter V. Marsden. Interpersonal Ties, Social Capital, and Employer Staffing Practices. In Nan Lin, Karen Cook, and Ronald S. Burt, editors, *Social Captial - Theory and Research*, pages 105–125. Aldine de Gruyter, New York, 2001.

James D Montgomery. *Social networks and persistent inequality in the labor market*. Center for Urban Affairs and Policy Research, 1991.

Perihan Saygin, Andrea Weber, and Michèle Weynandt. Coworkers, networks, and job search outcomes. 2014.

Rafael Polito Siano. Centralidade em redes de trabalho e diferenciais salariais entre ocupações no brasil. Technical report, Mestrado em Economia Aplicada -FEA-UFJF, 2016. URL https://repositorio.ufjf.br/jspui/handle/ufjf/3154.

Raymond T Sparrowe and Robert C Liden. Process and structure in leader-member exchange. *Academy of management Review*, 22(2):522–552, 1997.

Giorgio Topa. Social interactions, local spillovers and unemployment. *Review of economic studies*, pages 261–295, 2001.