**UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA**

The Role of Referral Hiring on Immigrant Labor Market Outcomes

O Papel da Contratação por Indicação no Mercado de Trabalho de Imigrantes

Marcos Ki Hyung Lee

Orientador: Prof. Dr. Danilo Camargo Iglori

**São Paulo - Brasil
2017**

Prof. Dr. Marco Antonio Zago
Reitor da Universidade de São Paulo

Prof. Dr. Adalberto Américo Fischmann
Diretor da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. Eduardo Amaral Haddad
Chefe do Departamento de Economia

Prof. Dr. Ariaster Baumgratz Chimeli
Coordenador do Programa de Pós-Graduação em Economia

MARCOS KI HYUNG LEE

The Role of Referral Hiring on Immigrant Labor Market Outcomes

O Papel da Contratação por Indicação no Mercado de Trabalho de Imigrantes

Dissertação apresentada ao Departamento de Economia da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo como requisito parcial para a obtenção do título de Mestre em Ciências.

Orientador: Prof. Dr. Danilo Camargo Iglioni

Versão Original

São Paulo - Brasil

2017

FICHA CATALOGRÁFICA

Elaborada pela Seção de Processamento Técnico do SBD/FEA/USP

Lee, Marcos Ki Hyung
The role of referral hiring in immigrant labor market outcomes /
Marcos Ki Hyung Lee. – São Paulo, 2017.
46 p.

Dissertação (Mestrado) – Universidade de São Paulo, 2017.
Orientador: Danilo Camargo Iglioni.

1. Mercado de trabalho 2. Salários 3. Imigrantes I. Universidade de
São Paulo. Faculdade de Economia, Administração e Contabilidade.
II. Título.

CDD – 331.12

Acknowledgements

This dissertation was only possible due to the support from my advisor, Professor Danilo Iglori, and the very helpful comments from Professor Renata Narita.

I thank the financial support from CNPq, which helped fund my studies, and the professors and researchers from NEREUS, which gave me access to the RAIS dataset and the computational infrastructure.

I also thank my classmates, who were always there to listen to all my grievances.

Finally, I am indebted to my family, which always encouraged me on my choices.

*“The truth will set you free.
But not until it is finished with you.”
David Foster Wallace*

Resumo

O objetivo dessa dissertação é investigar se a contratação por indicação leva a uma melhor busca por trabalho e melhores resultados no mercado de trabalho de imigrantes na cidade de São Paulo. Utilizamos uma *proxy* para a contratação por indicação, a proporção de trabalhadores da mesma nacionalidade um ano antes da contratação, e estimamos uma regressão de salários e demissões com a RAIS, um painel de firmas e trabalhadores formais. Nossos resultados mostram que a contratação por indicação diminui o salário do imigrante quando da contratação, isto é, se a *proxy* aumentasse em 10 p.p., salários cairiam em 1,9%. Também elaboramos alguns testes de placebo que reforçam nossos resultados. Identificamos que nacionalidades com trabalhadores menos educados são os que mais sofrem impacto negativo. Outros possíveis canais seriam o tamanho da rede de imigrantes e a idade de imigração das nacionalidades, mas não achamos nenhum impacto, apesar desses exercícios sofrerem de diversos problemas de identificação. Passos futuros seriam desenvolver um modelo teórico que racionalize os resultados.

Palavras-chave: mercado de trabalho, salários, imigrantes

Abstract

The aim of this dissertation is to investigate whether referral hiring leads to better job search and labor market outcomes for immigrants in São Paulo. To do so, we use a proxy for referral hiring, the share of same nationality workers in a firm one year before hiring, and estimate wage and turnover regression with a matched employer-employee dataset (RAIS). We find that referral hiring lower immigrant's wages at hiring, that is, if the proposed proxy increased by 10 p.p., wages would decrease by 1.9%. We also devise a couple of placebo test that reinforce the results we have found. We identify that nationalities with low education workers absorb all the negative impact. Other possible channels could be the size of immigrant networks and immigration age of nationalities, but we find no impact, although these exercises have several identification problems. Future steps include developing a theoretical framework that can rationalize the results.

Keywords: labor market, wages, immigrants

List of Tables

Table 1 – What was the last method you took to obtain a job?	19
Table 2 – Method of obtaining current job	19
Table 3 – Method of obtaining current job on education levels	20
Table 4 – Method of obtaining current job on wage levels	20
Table 5 – Method of obtaining current job on formality	20
Table 6 – Equation 3.2 Results	25
Table 7 – Equation 3.1 Results	26
Table 8 – Descriptive Statistics	27
Table 9 – Robustness Results	28
Table 10 – Nationality Hiring	29
Table 11 – Other Specifications	30
Table 12 – Correlations Between Network Sizes from Different Sources	31
Table 13 – Immigration Years from Census 2010	32
Table 14 – Other Specifications	33
Table 15 – Nationality Frequency	41
Table 16 – Nationalities by Immigration Age Groups	41
Table 17 – Nationalities by Network Size Groups	42
Table 18 – Nationalities by region	42
Table 19 – Regression Results	44

Contents

1	INTRODUCTION	15
2	DATA	19
3	EMPIRICAL STRATEGY	23
4	RESULTS	25
4.1	Main Results	25
4.2	Additional Results	30
5	CONCLUSION	35
	BIBLIOGRAPHY	37
	APPENDIX	39
	APPENDIX A – TABLES	41
	APPENDIX B – PME RESULTS	43

1 Introduction

Social interactions play a big role in economic interactions. They influence the decisions of economic agents like consumers and firms, making them relevant to understand economic problems. Ranging from peer effects in schools to financial networks, the applications are wide. The dynamic of these networks shape the outcomes of variables such as education attainment, crime behavior and labor markets. (TOPA, 2011; JACKSON, 2011; CALV6-ARMENGOL; PATAACCHINI; ZENOU, 2009)

Empirical and anecdotal evidence points out that some part of the current job search is done through informal social networks, such as friends, family and neighbors. This could potentially reduce search frictions and augment efficiency of job allocation. On the other hand, congestion and quality of networks might worsen labor market outcomes. Nevertheless, empirical evidence, specially for Brazil, is scarce.

It is also noticeable that immigrants and ethnic populations tend to form tight networks, due to several reasons (GROSS; SCHMITT, 2003). This fact usually implies that these groups rely more on referral methods than traditional ones, although not necessarily being more efficient (BATTU; SEAMAN; ZENOU, 2011; PATAACCHINI; ZENOU, 2012). With the recent immigration policy discussions in major economies, whether these populations have different job search dynamics and labor market outcomes than native populations is important to better direct policies towards them. Moreover, Sao Paulo city has had a long history of immigration, with really diverse groups coming since 1880, first with Italians, going through the Japanese and Lebanese immigration and more recently with Latin and Central American populations, among others throughout the recent history.

This dissertation aims to analyze the role of referral hiring in the São Paulo labor market for immigrant populations. More specifically, we are interested in whether referral hiring improves or not wages and turnover of immigrant workers in São Paulo. A straightforward attempt to understand this question would imply a OLS regression of wages/turnover on a variable indicating that this job was obtained through a referral. However, several identification issues arises, as the job search method is not used randomly across individuals and it is therefore correlated with observed and unobserved characteristics of the firm, the worker and the labor market itself.

We examine empirically this issue on the Sao Paulo labor market. First, we check if referral is a relevant job search method using data that has specific information about it. We find that referral is the second most used method of job search for workers (not just foreigners) in São Paulo, and that its use is more important for certain groups, such as informal and low educated workers.

Then, by relying on a matched worker-firm panel data to address identification problems, we use the share of same nationality workers in the firm one year before a foreign worker was hired as a proxy for a referral hiring. This proxy is based on an empirical strategy developed by [Dustmann et al. \(2016\)](#). Given fixed worker and firm fixed effects, if this share impacts wages and turnovers solely through referral, then we can interpret this proxy as probability of being hired through a referral. We see that this variable is a very good predictor of a firm's hiring behaviour by using a firm-nationality-year data and regressing the share of new hires of a specific nationality against our proxy.

We see that foreign workers that were hired with many same nationality workers in the firm one year before earn lower wages than if there were fewer same nationality workers. This is consistent with some placebo exercises. Moreover, this impact seems to be heterogenous on education levels. Nationalities with low education levels absorb all the negative impact, while high education nationalities show no impact whatsoever. Also, the hiring probabilities of nationalities with low educated worker are greater.

We interpret these results as referral being mostly important for low skilled immigrants. The PME gives us indirect evidence that they use more referrals as job search method. Our regressions shows us that they have a higher probability of obtaining jobs due to referral, but these jobs being low quality ones. This interpretation is compatible with the literature of a bad network "trap" ([CALVÓ-ARMENGOL; JACKSON, 2007](#); [BATTU; SEAMAN; ZENOU, 2011](#)).

One of the first works that attempted to model the relationship between social networks and labor market outcomes was [Montgomery \(1991\)](#). Building a two-period adverse selection model of the labor market with a simple social structure, the author assumes that firms will only ask for a referral to their high ability workers, thus producing a premium on wages of the referral market.

More recently, more complex works started to emerge. Perhaps the first study to model an explicit network in a dynamic framework was [Calvó-Armengol e Jackson \(2004\)](#). According to them, agents receive information from job openings at a certain probability. Those who are already employed will randomly pass this information to a person from his network that is unemployed. Because their model is dynamic, it allows them to conclude that, even though on the first moment being in a network with a lot of unemployed agents may be harmful because of competition, on the long run the extent of the network has a multiplier effect, because friends of friends of friends and etc will affect their own status.

In a generalization of their first model, [Calvó-Armengol e Jackson \(2007\)](#) includes wage dynamics, and show that making ones neighbors better off or adding new agents to the network will improve the agent's future wages. Also, adding to the literature of racial wage inequality, the authors show that if two different social groups with the same information network start out with different wage and employment levels, there will be

persistent differences of drop-out rates and wages in the long run. This highlights the fact that, although the two groups have the same job information network, the state of the agent's network have important long-run effects.

Tying to the job search-and-match literature, works like [Fontaine \(2008\)](#) and [Dustmann et al. \(2016\)](#) have incorporated the role of social networks into models with on-the-job search and wage bargaining. For instance, the first author shows that wage dispersion arises in a context of unequal arrival rates, which are caused by different unemployment rates across networks. On the other hand, the latter models networks in a way that lowers the noise of the worker's real productivity, so referral hires are on average better than others, with higher wages and lower turnovers.

The empirical body of work has approached the referral effect in two ways. The way most of them have proceeded is to approximate social space to geographical space. It can be reasonable to consider that attributes not only define social space such as ethnicity, race, age, etc, but also interact with physical distance, since the cost to maintain an interaction increases with it. Although arguably the people in your friends and relatives' network do not necessarily live in your neighborhood, it is a simplified hypothesis due to modeling and data constraints. Specifically in labor market themes, detailed data of networks is hard to come by ([TOPA; ZENOU, 2015](#)).

As noted by [Manski \(1993\)](#), identification of social network effects on labor market outcomes suffer from what is called the reflection problem, i.e., that it is difficult to distinguish individual decisions that are potentially affected by aspects from the network from the characteristics of the network, which in turn impact the individual's decision. Careful econometric strategy has to be employed to address this.

Earlier works on this theme used to focus on natural experiments evidence. Both the Moving to Opportunity program, which randomly assigned housing vouchers for volunteers, or the resettlement of refugees in various cities have served as experiments ([TOPA; ZENOU, 2015](#)).

[Bayer, Ross e Topa \(2008\)](#) developed a novel identification strategy that do not depend on natural experiments, and it was used in later studies by other authors. Noticing the difficulties of properly identifying neighborhood effects because of the sorting and unobserved heterogeneity of individuals and neighborhoods, the authors use Census data from the Boston metropolitan area which details the city block where individuals live and work to compare outcomes for neighbors who reside in the same block against those who live in the same neighborhood, but not on the same block. The findings are that the probability of working together if residing on the same block is 33% greater than if residing on nearby blocks. The underlying identification hypothesis is that, given the neighborhood, there is no correlation in unobserved traits at the block level.

Another way to measure the importance of referrals is to comprehend what impacts it has on the duration of job spells and wage levels of jobs obtained through referrals. [Dustmann et al. \(2016\)](#), for example, shows that indeed job search networks improve labor market outcomes by reducing informational frictions and improving productivity. They use an employer-employee matched panel data to test whether worker's wages and layoff rates are better if they obtained their job through a referral, and their evidence points to that direction.

In order to investigate social networks and their effect on wage premiums, [Schmutte \(2015\)](#) builds a sample of job-to-job movers from 2002 to 2003 and a variable of each workers' job quality information network that consists of the average wage premium of all the employed workers that live in his own block. Their main finding is that neighbors have a high probability of working together and of becoming employed in high paying jobs when their neighbors already work in high wage firms.

Focusing on ethnic minorities, [Patacchini e Zenou \(2012\)](#) explores the relationship between finding a job through social networks and the residential proximity of ethnic minorities. Following one of the implications of [Calvó-Armengol e Jackson \(2004\)](#) that the length of weak ties lowers the probability of finding jobs, their result is that this probability for minorities decreases rapidly with physical distance. Also, living with large numbers of employed neighbors of the same ethnicity raises the likelihood of having jobs.

On the other hand, [Battu, Seaman e Zenou \(2011\)](#) argues that ethnic networks do not necessarily help recent immigrants, for example, and that the longer these individuals reside in the UK, the more they resort to direct methods of finding jobs.

Contributing to the literature on the "adverse" effects of networks, [Battisti, Peri e Romiti \(2016\)](#) develops a model with two job search methods, a formal one and a network-referral based one. Their findings show that the size of the ethnic network upon arrival of the individual in the country are positively correlated with employment but negatively correlated with human capital. Also, in the long run, wages and match quality are worse for them, compared to immigrants that had a smaller network size. This is related to the conclusion of [Calvó-Armengol e Jackson \(2007\)](#) that networks where the individuals have low quality of employment have persistent long run effects.

2 Data

Some theoretic and empirical works have found that jobs obtained through referral have better outcomes in terms of wages, turnover, duration, etc. But firstly, it is important to establish if referral hiring is a relevant characteristic of the Brazilian and the São Paulo labor market.

The PME (Monthly Employment Survey, in Portuguese) allows us to try to find some evidence of this matter in the São Paulo labor market. It is a longitudinal rotational survey that interviews households for seven Metropolitan Areas that can produce statistics about fluctuations and tendencies of the urban labor market. One of the questions of the survey asks what was the last method one took to find a job for all workers who never worked or that left his previous job at least one year before.

Table 1 – What was the last method you took to obtain a job?

	Percentage
Contacted Firm Directly	77.67%
Referred to Relatives or Friends	8.30%
Looked Into Job Agencies	7.92%
Others	6.11%

Source: PME 2007-2012. Other methods include public: took exam for public sector jobs, consulted ads, consulted worker unions and took initiative to open own business.

As can be seen on Table 1, although only 8.30% answered that they referred to relatives or friends, it is still the second most relevant method. Also, restricting to those that were asked and got a job in the subsequent interview, we see that referral still maintains its relative importance in Table 2.

Table 2 – Method of obtaining current job

	Percentage
Contacted Firm Directly	73.30%
Referred to Relatives or Friends	11.60%
Looked Into Job Agencies	7.88%
Others	7.22%

Source: PME 2007-2012. Other methods include public: took exam for public sector jobs, consulted ads, consulted worker unions and took initiative to open own business.

Table 3 – Method of obtaining current job on education levels

	Low Education	High Education
Contacted Firm Directly	74.54%	72.20%
Referred to Relatives or Friends	17.13%	10.79%
Looked Into Job Agencies	4.63%	6.64%
Others	3.70%	10.37%

Source: PME 2007-2012. Low education is defined as up to 10 years of years of study. Other methods include public: took exam for public sector jobs, consulted ads, consulted worker unions and took initiative to open own business. Largest contribution to change in “Others” category are public sector exams and initiative to open own business.

Table 4 – Method of obtaining current job on wage levels

	Low Wage	High Wage
Contacted Firm Directly	74.89%	71.56%
Referred to Relatives or Friends	14.29%	9.17%
Looked Into Job Agencies	6.49%	9.17%
Others	4.33%	10.10%

Source: PME 2007-2012. Low wage is defined as below the sample median. Wage is for the job obtained the month after question was answered. Other methods include public: took exam for public sector jobs, consulted ads, consulted worker unions and took initiative to open own business. Largest contribution to change in “Others” category are public sector exams and initiative to open own business.

Table 5 – Method of obtaining current job on formality

	Informal	Formal
Contacted Firm Directly	76.23%	72.26%
Referred to Relatives or Friends	13.00%	9.68%
Looked Into Job Agencies	7.62%	8.39%
Others	3.15%	9.67%

Source: PME 2007-2012. Formality is defined for the job obtained the month after question was answered. Other methods include public: took exam for public sector jobs, consulted ads, consulted worker unions and took initiative to open own business. Largest contribution to change in “Others” category are public sector exams and initiative to open own business.

Conditioning this question to education and wage levels¹, we see on Tables 3 and 4 that referral raises in importance for the low education and low wage individuals. Also, in Table 5 we see that referral is more important in obtaining informal jobs. This seems to indicate that referral is more relevant to certain groups than others.

Since the PME has a direct variable of referral, a straightforward exercise would be

¹ The wage being considered is the one from the job the individual obtained.

to analyze its impact on labor market outcomes. Nevertheless, we can't control for several omitted variables, such as worker and firm fixed effects.² Because of this limitation, we consider that the RAIS, a matched worker-employer panel survey available from 2003 to 2013, is more adequate. It is a nation-wide survey that has to be answered by all formal firms every year, with information about firm and worker characteristics. Besides key variables such as occupation, wages, education levels and work history, we also have worker nationality, which we use to define immigrants. Therefore, naturalized and ethnic but brazilian born workers are not considered to be immigrants. We restrict ourselves to the city of Sao Paulo, because it has had a very long history with immigration and has a very dynamic labor market.

² In the Appendix, some correlation regressions can be found. Its main results are that the probability of getting a job through referral is on average 12% higher than the other methods.

3 Empirical Strategy

With this framework, it is possible to apply the methodology developed by [Dustmann et al. \(2016\)](#) to São Paulo. Following their notation, we would like to estimate:

$$y_{ijt} = \beta_0 + \beta_1 Ref_{ij}^\tau + \beta_2 Ref_{ij}^\tau \cdot tenure_{ijt} + X'_{ijt}\beta_3 + \gamma_t + \delta_i + f_j + \epsilon_{ijt} \quad (3.1)$$

Where y_{ijt} can be the log wage of worker i or an indicator equal to 1 if the worker leaves the firm in $t + 1$, Ref_{ij}^τ is equal to 1 if the worker obtained the job through referral on firm j in year τ , X_{ijt} is a vector of time varying characteristics and γ_t , δ_i and f_j are year, worker and firm fixed effects, while ϵ_{ijt} is the error term. If referral has a positive impact on wages and turnover, we should expect that $\beta_1 > 0$ and $\beta_2 < 0$ and $\beta_1 < 0$ and $\beta_2 > 0$ for the wage and turnover regression respectively (because firms and workers learn over time).

Both the database used by the authors and the RAIS do not offer information about referral hiring, so they use a proxy variable motivated by their theoretical model: the share of workers from the same minority group in the firm one period before the worker was hired, $S_{gj}^{\tau-1}$, where g is the minority group of the individual. If immigrant workers refer people of their own ethnicity to their employer, then the higher the share of foreign workers in a firm, the more likely it is that the firm will hire workers from group g through a referral.

Another important assumption for the proxy variable to work is that conditional on worker and firm fixed effects, this share of workers from the same minority group in the firm one period before the worker's hiring must impact wages and turnovers only through referrals. We can argue that this is an acceptable assumption for several reasons. First, worker fixed effects control for the fact that certain nationalities or type of workers may tend to use more referral as a means of job search (as shown in the PME survey). Second, the inclusion of firm fixed effects take into account the possibility that certain types of firms might use more referral, or hire more from a specific nationality. Lastly, both inclusions control for non-random sorting of workers and firms, that is, low-skilled workers sort into firms that employ workers from the same nationalities.

Due to the size of the database and the number of fixed effects, when estimating equation 3.1 we used only the years of 2009 to 2013 and employed an algorithm developed by [Correia \(2014\)](#) to calculate high dimension fixed effects and clustered standard errors.¹ Also, we interact all variables of equation 3.1 with a foreign dummy, so all coefficients

¹ See [Cameron e Miller \(2015\)](#) for a discussion of clustering and fixed effects models.

are relative to only foreign workers. We do not simply drop all brazilian workers because they are important when calculating firm fixed effects.² To ensure that we know the whole history of firm workers, we also drop everyone that started working before 2003.³ The share of immigrant workers in the firm is calculated using all the years available, from 2003 to 2013, before any restriction is imposed, and use only from 2009 onwards.⁴

To motivate this, if the own group share is positively related to the probability of obtaining a job through a referral, then firms that employed a higher share of minority workers in the past will be more likely to hire from that group in the future. To test this, the following regression will be estimated:

$$H_{gj}^{\tau} = \alpha_0 + \alpha_1 S_{gj}^{\tau-1} + X_j^{\tau} \alpha_{2g} + Z_{gj}^{\tau-1} \alpha_3 + \gamma_g^{\tau} + u_{gj}^{\tau} \quad (3.2)$$

where H_{gj}^{τ} is the share of minority workers from group g among all new hires in firm j at time τ , $S_{gj}^{\tau-1}$ is the share of minority workers from the same group in the firm in $\tau - 1$, X_j^{τ} and $Z_{gj}^{\tau-1}$ are demand and supply side controls, γ_g^{τ} is minority group-specific year fixed effects and u_{gj}^{τ} is the error term. If α_1 is positive and significant, then there is evidence that the proposed proxy variable is valid, that is, that the share of same nationality workers in $t - 1$ is a good predictor for the firm's hiring behaviour. Also, it can also be a good indicator that foreigner workers are benefited in terms of a greater probability of finding a job due to the presence of same nationality workers. We do not use firm fixed effects in this case because then we would be eliminating most of the variation, and the fact that firms with many foreigners tend (or not) to hire more foreigners of the same nationality would be entirely captured by the fixed effects.

² To speed up the estimation, we drop all firms that never had or hired a foreign worker. This does not change the estimates.

³ If we did not do this, then there would be foreign workers which we did not know his share when he was hired.

⁴ We do this for two reasons, one already mentioned. It speeds up estimation and we also allow for more variation in the data regarding tenure, since we only have workers that started working after 2003.

4 Results

4.1 Main Results

Results from equation 3.2 can be found in Table 6. To estimate this model, we built a firm-nationality-year database from 2003 to 2013, using the 17 largest minorities. Results indicate that the share of foreigners in $t - 1$ helps predict the firm's hiring behaviour. If the share from a nationality in a firm before hiring increases by 10 percentage points, the share of foreigners from that group among the new hires will increase by 4.2 percentage points. Also, this parameter can be interpreted as the probability of obtaining a job through referral¹, meaning that 42% of foreigners obtain jobs through referral.^{2 3}

Table 6 – Equation 3.2 Results

	(1)	(2)	(3)
S_{gj}^{t-1}	0.371*** (0.028)	0.417*** (0.032)	0.414*** (0.032)
Control	No	Yes	Yes
Observations	31,343,568	26,992,848	26,992,848
Firms	486,397	441,033	441,033
R-squared	0.997	0.9995	0.9995

Standard errors are clustered in firms. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Nationality-year specific fixed effects included. Dependent variable is the share of foreigners of a nationality over all new hires in t . Only firms that hired someone in a year were used. The largest 17 minority groups were used, which make up almost 95% of all foreigners. Model 2 includes share of women, low education and mid education new hires as controls, which are all interacted with nationality dummies. Model 3 includes the share of workers from other nationalities as a proxy for the firm's openness to foreigners. All observations are weighted by the number of new hires in a firm on a given year because data are group-mean. When not weighted, estimates drop to around 0.25 at 1% significance.

This result indicates then that foreign workers are benefited by a large presence of same nationality workers in a firm through a higher probability of being hired.

Looking at the impact of the share of same nationality workers one year before the hiring on wages on Table 7, it seems that increasing the share of workers from the same nationality one year before hiring by 10 p.p. would decrease their wages by 1,9%.

This result is consistent throughout most specifications, including the most robust one, with worker and firm fixed effects. For turnover decisions, results are inconsistent,

¹ This is implied by the model developed by [Dustmann et al. \(2016\)](#)

² The raw number from the PME is lower, around 12%, but for the whole population and not only foreigners.

³ These results are similar to the ones found by [Dustmann et al. \(2016\)](#)

with the final specification not showing results, although the point estimate might indicate a lower probability of leaving the firm when there are same-nationality workers on the firm.

Table 7 – Equation 3.1 Results

		(1) OLS	(2) OLS, Controls	(3) Worker Fixed Effects	(4) Firm Fixed Effects	(5) Worker & Firm Fixed Effects
Baseline						
Wages	$S_{gj}^{\tau-1}$	-1.374*** (0.019)	-0.119*** (0.018)	-0.104*** (0.021)	0.034 (0.157)	-0.187** (0.034)
Turnover	$S_{gj}^{\tau-1}$	0.223*** (0.009)	0.142*** (0.010)	0.173*** (0.034)	0.136*** (0.025)	-0.078 (0.055)
Tenure Interaction						
Wages	$S_{gj}^{\tau-1}$	-2.525*** (0.146)	-0.253** (0.105)	-0.038 (0.045)	-0.045 (0.114)	-0.001 (0.048)
	$S_{gj}^{\tau-1} \times tenure$	1.069*** (0.145)	0.123*** (0.104)	-0.076** (0.037)	0.051 (0.168)	-0.069** (0.036)
Turnover	$S_{gj}^{\tau-1}$	0.547*** (0.025)	0.204*** (0.028)	-0.002*** (0.060)	0.157 (0.058)	-0.190* (0.098)
	$S_{gj}^{\tau-1} \times tenure$	-0.337*** (0.021)	-0.068*** (0.024)	0.201*** (0.047)	-0.009 (0.030)	0.229** (0.074)

Standard errors are clustered in workers and on fixed effects. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Sample from 2009 to 2013. To save computation time, an algorithm developed by [Correia \(2014\)](#) was used, which drops some observations which are not used when calculating fixed effects. The control variables are: tenure categories, log of firm size, age, age squared, firm sector dummies, education and gender dummies. All variables are interacted with a dummy for foreign status. Brazilian workers are pooled together so that their fixed effects are also estimated. Number of observations vary from 28,419,862 to 25,907,714 in the wage regression and from 28,874,618 to 26,360,178

The second equation interacts our variable of interest with tenure, to check if the results found above increase or decrease over time. For wages, the results seem to reinforce the fact that these workers have lower wages, even after some time in the firm. For turnover, the previously expected results appears, indicating that these workers have a lower probability of leaving, but that it decreases over time, possibly because of learning within the firm. While [Dustmann et al. \(2016\)](#) methodology confirms that minority workers in Munich have higher wages and lower turnover rates when hired through referral, when applied to Sao Paulo, the results on wages are the opposite, seemingly supporting models that predict negative effects of networks on foreigners.⁴

Since the RAIS dataset only captures formal workers, these results could be suffering from a selection bias. We argue that this selections bias is at least underestimating the

⁴ [Dustmann et al. \(2016\)](#) restricts the sample to only low and medium-skilled workers due to wage censoring on their data, which could partly explain these results.

effect on wages and turnovers. Using the PME survey, we have shown that low education workers or those that obtained informal jobs or jobs that pay lower wages tend to use more referral methods. If this holds up for foreigner workers too, and we consider that informal workers also tend to receive lower wages, then our regression is underestimating the effect due to the absence of this fraction of foreign workers in our dataset. Hence, the negative effects on wages and turnovers could be potentially greater if taken into account informal workers.

This contradicting result regarding wages and turnover might be related to the heterogeneity of the different ethnic groups in Sao Paulo, that is, their networks could have different qualities. Whereas groups from OECD countries tend to have high education levels and relatives with high quality jobs in Sao Paulo, immigrants from poorer countries form networks with low employment levels, for example. This heterogeneity, considering that education levels didn't change over the years for these workers, is controlled by fixed effects, so they do not explicitly show up on the results.

Table 15 in the Appendix separates nationalities by their median education level.⁵ It is noticeable that the low and high education groups are each composed by nationalities of the same region, and that some anecdotal patterns emerge, such as the East Asian minorities belonging to the high education one. In Table 8, descriptive statistics are shown for the low and high education.

Table 8 – Descriptive Statistics

	All				Foreigners				High Educ Foreigners				Low Educ Foreigners			
	Mean	Median	Min	Max	Mean	Median	Min	Max	Mean	Median	Min	Max	Mean	Median	Min	Max
Women	0.43	0	0	1	0.32	0	0	1	0.33	0	0	1	0.31	0	0	1
Turnover	0.41	0	0	1	0.31	0	0	1	0.27	0	0	1	0.38	0	0	1
Wage	7.16	6.97	3.50	11.53	8.05	7.69	4.86	11.52	8.44	8.38	4.86	11.52	7.18	6.96	5.06	11.34
Firm Size	5.35	5.23	0.69	10.73	4.62	4.42	0.69	10.73	4.86	4.86	0.69	10.73	4.10	3.66	0.69	10.73
Network Size	15.55	15.60	1.39	15.66	7.12	7.24	1.39	8.69	6.84	6.94	2.30	7.91	7.74	8.07	1.39	8.69
$S_{ij}^{\tau-1}$	0.888	1	0	1	0.049	0	0	1	0.025	0	0	1	0.104	0	0	1

First group has all nationalities, including Brazilians. Low and High Education groups were separated by their median education level. RAIS dataset from 2009 to 2013. Wages, Firm Size and Network size are in log. Network size is the total number of workers of that nationality.

To understand whether these characteristics are important to characterize referral effects, two models were estimated, one with only Low Education nationalities and another with only High Education ones. Results are in Table 9. On the baseline model, wages are negatively affected by the share of same nationality workers only in the low education group, while all other coefficients are not statistically significant. When adding the interaction with tenure, there is no significant effect in column (2), but expected positive impacts of referral show up in column (1) for turnover.

⁵ Fortunately, nationalities were separated exactly in two groups, Secondary Education and Superior Education workers. As such, these were used to classify Low and High education groups.

Table 9 – Robustness Results

		(1)	(2)
		Low Education	High Education
Baseline			
Wages			
	$S_{gj}^{\tau-1}$	-0.063** (0.028)	-0.132 (0.090)
Turnover			
	$S_{gj}^{\tau-1}$	0.008 (0.063)	0.110 (0.098)
Interactions			
Wages			
	$S_{gj}^{\tau-1}$	0.112** (0.059)	-0.111 (0.111)
Turnover	$S_{gj}^{\tau-1} \times tenure$	-0.171*** (0.050)	-0.012 (0.050)
	$S_{gj}^{\tau-1}$	-0.400*** (0.129)	0.078 (0.149)
	$S_{gj}^{\tau-1} \times tenure$	0.389*** (0.108)	0.135 (0.092)

Standard errors are clustered in workers and on fixed effects. Worker and Fixed effects included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Sample from 2009 to 2013. To save computation time, an algorithm developed by [Correia \(2014\)](#) was used, which drops some observations which are not used when calculating fixed effects. Low and High Education groups were separated by their median education level. Network size (in log) is the total number of workers of that nationality each year. All variables are interacted with a dummy for foreign status. Brazilian workers are pooled together so that their fixed effects are also estimated. Number of observations vary from 28,419,862 to 25,907,714 in the wage regression and from 28,874,618 to 26,360,178

To further understand the heterogeneity of the referral effect on immigrants, now we estimate hiring regressions with country specific dummies. In this way, we can assess if certain nationalities suffer stronger impacts. We consider that equation 3.2, which was devised as a test for the validity for the proxy variable, also captures the hiring behaviour and therefore the impact on hiring. For this regression, we interacted our main variable with a dummy for all 17 nationalities used.⁶ The same was not done with the wage and turnover regressions due to the very low variability in the data for many nationalities when including firm and worker fixed effects.

Table 10 shows that, in general, nationalities in the low education group tend to have a higher probability of being hired, besides the previous result that they also have the greater negative impact on wages.

The results for this robustness exercise could be interpreted as referral being important only to low skilled immigrants. This is consistent with the patterns found in

⁶ Again, we restrict ourselves to these nationalities due to estimation speed restrictions, and the fact that they encompass 95% of all foreigners

Table 10 – Nationality Hiring

Nation	Secondary Education	Nation	Superior Education
	Hiring		Hiring
Bolivia	0.779*** (0.027)	Portugal	0.061*** (0.01)
Paraguay	0.51*** (0.064)	Argentina	0.122*** (0.03)
Haiti	0.447*** (0.146)	Chile	0.048*** (0.009)
Peru	0.409*** (0.095)	China	0.577*** (0.153)
		USA	0.486*** (0.184)
		Japan	0.349*** (0.067)
		Spain	0.146*** (0.041)
		France	0.135** (0.06)
		Italy	0.098*** (0.035)
		Uruguay	0.02*** (0.006)
		Germany	0.265*** (0.045)
		South Korea	0.285*** (0.039)
		UK	0.356*** (0.13)

Standard errors are clustered in firms. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Coefficients from estimates of equation 3.2 where the proxy variable was interacted with nationality dummies. Nationality-year specific fixed effects included. Only firms that hired someone in a year were used. The largest 17 minority groups were used, which make up almost 95% of all foreigners. All observations are weighted by the number of new hires in a firm on a given year because data are group-mean.

the PME survey, where individuals with low income and low education depended more on referrals. Since low educated foreigners need to rely more on referrals between their counterparts, which are itself low skilled and in worse jobs, it creates a sort of “trap” in a bad network, a hypothesis found in [Calvó-Armengol e Jackson \(2007\)](#) and [Battu, Seaman e Zenou \(2011\)](#). Nevertheless, these workers are still getting jobs, perhaps more so than if they did not rely on their networks, which is what we observe above. Also, considering the theoretical model by [Dustmann et al. \(2016\)](#), where referral reduces the noise firms face when observing a worker’s true productivity, it seems that for high skilled immigrants this noise is less important, maybe because they have other ways to decrease it, like past experience.

4.2 Additional Results

To make sure we are identifying the effect we are seeking, we devised a couple of placebo tests. First, to understand whether we are capturing a spurious coefficient or not, we keep only foreign workers that were hired when there were no other same nationality workers in $t - 1$ and build a new variable that is the share of all other foreigners in the firm in $t - 1$. If the impact that we captured in our main results is true, we should expect that there would be no effect, since the channel through which workers gain/lose wages or turnover rates is solely the presence of same nationality coworkers in $t - 1$. On the other hand, if we just keep firms that hired workers of a certain nationality in which in $t - 1$ there were only workers of that same nationality (excluding brazilians), then our main impact would be intensified. Results are in Table 11.

Table 11 – Other Specifications

		(1) Baseline	(2) Other Nationalities	(3) Own Nationality
Set to zero				
Wages				
	$S_{gj}^{\tau-1}$	-0.075** (0.030)	0.132 (0.113)	-0.163*** (0.034)
Turnover				
	$S_{gj}^{\tau-1}$	0.056 (0.054)	-0.567*** (0.124)	-0.111* (0.061)
Dropped				
Wages				
	$S_{gj}^{\tau-1}$	-0.075** (0.030)	0.146 (0.132)	-0.193** (0.041)
Turnover				
	$S_{gj}^{\tau-1}$	0.056 (0.054)	-0.561*** (0.149)	-0.157** (0.064)

Standard errors are clustered in fixed effects. Worker and Fixed effects included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Sample from 2009 to 2013. To save computation time, an algorithm developed by [Correia \(2014\)](#) was used, which drops singleton observations which are not used when calculating fixed effects. Column 1 shows the baseline results. Column 2 keeps only workers that were hired in which the firm in $t-1$ had no other own nationality workers, but had other nationality workers present. On the other hand, column 3 keeps only workers that were hired in which the firm in $t-1$ had only workers from the same nationality, excluding cases in which there were other nationalities present. All variables are interacted with a dummy for foreign status. Brazilian workers are pooled together so that their fixed effects are also estimated. Number of observations vary from 28,419,862 to 25,907,714 in the wage regression and from 28,874,618 to 26,360,178

Results from the placebo tests for the wage regression follow our expectations, in which we find no effect for the first placebo and a greater effect in the same direction for the second placebo. In the turnover regression, in both cases we get a negative and statistically significant impact. This seems to indicate that the mere presence of foreigners in your firm one year before being hired is enough to lower your turnover probability. Conversely, it might be just a spurious correlation.

Another important characterization of immigrant groups are their size. Newly arrived immigrants could benefit from a large sized network due to the greater information available or they could be constrained by it because they are limited to the referral channels and do not employ active search through other methods. Likewise, years since immigration can be relevant to determine wages and turnovers of foreigners as noted by [Battu, Seaman e Zenou \(2011\)](#) and [Battisti, Peri e Romiti \(2016\)](#). On the one hand, nationalities that have arrived some time ago, and thus have been more or less assimilated to the local labor market, should use more formal methods of job search, and the impact of having friends in a firm of your own nationality would matter less. On the other hand, if these groups are larger, there could be some substantial congestion effect, generating a more negative impact on both wages and turnovers.

Since the RAIS does not offer any information on immigration years, it is not possible to do this analysis. However, the Census Survey for 2010 does offer this information on a micro level. It asks every foreigner surveyed which year they arrived in Brazil. With this information, it is possible to assign a median value of years of immigration for every nationality in Sao Paulo, and then relate it to our dataset. Network size can be calculated using the frequency of the RAIS data or the Census data. We expect that network size and immigration years are negatively correlated, because older immigration groups are already in their second, third or more generation. Both measures have a high correlation coefficient, as can be seen in Table 12, and immigration years and nation frequency from RAIS have the expected signal.

Table 12 – Correlations Between Network Sizes from Different Sources

	Immigration Years	Nat Freq (RAIS)	Nat Freq (Census)
Immigration Years	1		
Nation Frequency (RAIS)	-0.2170	1	
Nation Frequency (Census)	0.3803	0.6824	1

Note: Immigration Years were calculated using the Census Survey from 2010. All variables are at nationality level.

Table 13 shows the first and last five nationalities in our dataset in terms of their median immigration years from the Census Survey.⁷

We create categories of “degree of assimilation” and of network size and interact them both with our main variable of interest.^{8 9} With this new regression, we expect that

⁷ It is important to note that the Haitian mass immigration to Brazil started in 2010, so the Census Survey does not capture that information, having only one occurrence of a Haitian immigrant that arrived in 19XX, while the RAIS dataset already incorporates this phenomenon. This information was manually changed to reflect it.

⁸ Because it is a time-invariant variable, we can only do a simple interaction, that is, if x is our variable of interest and y is our category variable, interactions such as $x + y + x \times y$ are not possible, only $x \times y$. Thus, we are essentially doing four different regression, one for each category.

⁹ Table 16 and 17 in the Appendix shows what nations are in each group.

Table 13 – Immigration Years from Census 2010

Top 5		Bottom 5	
Nation	Immigration Years (Median)	Nation	Immigration Years (Median)
Italy	56	Venezuela	6
Spain	54	Colombia	5
Portugal	53	Paraguay	4
Japan	52	Ecuador	2
Germany	50	Haiti	2

Note: Immigration Years are median values from the Census Survey from 2010.

groups of longer immigration history and greater network size presents either similar or more negative coefficients.

In Table 14, we see that estimates for the wage model are negative overall for all quartiles of network sizes and immigration ages. For turnover, we have no significant in general, similar to the baseline results. This probably indicates that this categorization was more of a random grouping than one with some economic significance. Nevertheless, these two measures proposed suffer from identification and measurement errors, which could be explaining their seemingly absence of economic relevance.

Table 14 – Other Specifications

		(1)	(2)
	Categories	Wages	Turnover
Immigration Years			
	(P0-P25)	-0.217*** (0.067)	-0.009 (0.105)
	(P25-P50)	-0.145*** (0.038)	-0.091 (0.066)
$S_{gj}^{\tau-1} \times$	(P50-P75)	-0.666*** (0.185)	-0.214 (0.197)
	(P75-P100)	-0.273* (0.273)	-0.086 (0.202)
Network Size			
	(P0-P25)	-0.335** (0.160)	-0.462** (0.203)
	(P25-P50)	-0.226*** (0.068)	0.022 (0.085)
$S_{gj}^{\tau-1} \times$	(P50-P75)	-1.145*** (0.236)	-0.408 (0.288)
	(P75-P100)	-0.124*** (0.035)	-0.085 (0.071)

Standard errors are clustered in fixed effects. Worker and Fixed effects included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Sample from 2009 to 2013. To save computation time, an algorithm developed by [Correia \(2014\)](#) was used, which drops singleton observations which are not used when calculating fixed effects. Column 1 shows the results for wages, while Column 2 for turnovers. Immigration Years calculated from Census 2010. Network size calculated from frequency of nationalities in RAIS. All variables are interacted with a dummy for foreign status. Brazilian workers are pooled together so that their fixed effects are also estimated. Number of observations vary from 28,419,862 to 25,907,714 in the wage regression and from 28,874,618 to 26,360,178

5 Conclusion

Referral hiring is a very important component of the job search in the Sao Paulo labor market. It is also a city that has always had a rich immigration history. In this dissertation, we examine the relation between referral hiring and immigrant labor market outcomes in Sao Paulo. To do so, we use a matched employer-employee panel dataset of all formal workers and consider the impact of a proxy of referral hiring on wages and turnover.

The proxy we use to identify a referral hiring for a specific nationality is the share of that nationality in the total of workers in a firm one year before the hiring. We test the validity of this proxy by regressing the share of that nationality in all new hires against it, and we see that it is indeed a good predictor for the firm hiring behaviour, with an increase in 10 p.p. of the proxy increasing the share of foreigners of that group among new hires by 4.2 percentage points. Also, this predictor is stronger for certain nationalities, in particular ones with workers with low education.

Moreover, the impact of that proxy, and thus referral hiring, on wages is negative while for turnover it has no statistical significance. That is, if the share of workers of a nationality increased by 10 p.p. one year before the hiring, then wages would decrease by 1.9%. Identification issues are dealt by including relevant sets of controls and firm and fixed effects, given that conditional on them, the proxy only impacts wages and turnovers through referral.

These results are consistent with some of the literature of the adverse effects of referral. Although these low skilled workers are getting more jobs, they are not high quality ones, due to the fact that these workers rely more on referral and thus are trapped in bad networks.

This effect on wages is heterogeneous on education, where nationalities with low education workers keep the negative impact, while high education ones show no impact at all. This is consistent with data from the PME that these groups, on a general level and not restricted to foreigners, tend to use more referral hiring. Placebo tests reinforce the validity of these results.

Finally, we devise an exercise to check the heterogeneity of these impacts on network sizes, immigration age of nationalities and region or country of origin. They are mostly inconclusive, but also suffer from a great deal of measurement errors and selection bias. Further steps would include a formal model to motivate the findings.

Bibliography

BATTISTI, M.; PERI, G.; ROMITI, A. *Dynamic Effects of Co-Ethnic Networks on Immigrants' Economic Success*. [S.l.], 2016. Citado 2 vezes nas páginas 18 and 31.

BATTU, H.; SEAMAN, P.; ZENOU, Y. Job contact networks and the ethnic minorities. *Labour Economics*, v. 18, n. 1, p. 48–56, 2011. Citado 5 vezes nas páginas 15, 16, 18, 29, and 31.

BAYER, P.; ROSS, S. L.; TOPA, G. Place of work and place of residence: Informal hiring networks and labor market outcomes. *Journal of Political Economy*, v. 116, n. 6, p. 1150–1196, 2008. Citado na página 17.

CALVó-ARMENGOL, A.; JACKSON, M. O. The effects of social networks on employment and inequality. *American Economic Review*, v. 94, n. 3, p. 426–454, 2004. Citado 2 vezes nas páginas 16 and 18.

CALVó-ARMENGOL, A.; JACKSON, M. O. Networks in labor markets: Wage and employment dynamics and inequality. *Journal of Economic Theory*, v. 132, n. 1, p. 27–46, 2007. Citado 3 vezes nas páginas 16, 18, and 29.

CALVó-ARMENGOL, A.; PATAcCHINI, E.; ZENOU, Y. Peer effects and social networks in education. *The Review of Economic Studies*, v. 76, n. 4, p. 1239–1267, 2009. Citado na página 15.

CAMERON, A. C.; MILLER, D. L. A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, University of Wisconsin Press, v. 50, n. 2, p. 317–372, 2015. Citado na página 23.

CORREIA, S. *REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects*. 2014. Statistical Software Components, Boston College Department of Economics. Disponível em: <<https://ideas.repec.org/c/boc/bocode/s457874.html>>. Citado 5 vezes nas páginas 23, 26, 28, 30, and 33.

DUSTMANN, C. et al. Referral-based job search networks. *Review of Economic Studies*, p. 1 – 33, 2016. Citado 7 vezes nas páginas 16, 17, 18, 23, 25, 26, and 29.

FONTAINE, F. Why are similar works paid differently? the role of social networks. *Journal of Economic Dynamics & Control*, v. 32, p. 3960–3977, 2008. Citado na página 17.

GROSS, D. M.; SCHMITT, N. The role of cultural clustering in attracting new immigrants. *Journal of Regional Science*, v. 43, n. 2, p. 295–318, 2003. Citado na página 15.

JACKSON, M. O. Chapter 12 - an overview of social networks and economic applications*. In: BENHABIB, A. B. J.; JACKSON, M. O. (Ed.). [S.l.]: North-Holland, 2011, (Handbook of Social Economics, v. 1). p. 511 – 585. Citado na página 15.

MANSKI, C. F. Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, v. 60, p. 531–542, 1993. Citado na página 17.

MONTGOMERY, J. D. Social networks and labor-market outcomes: Toward an economic analysis. *American Economic Review*, v. 81, n. 5, p. 1408–1418, 1991. Citado na página 16.

PATACCHINI, E.; ZENOU, Y. Ethnic networks and employment outcomes. *Regional Science and Urban Economics*, v. 42, p. 938–949, 2012. Citado 2 vezes nas páginas 15 and 18.

RIBAS, R. P.; SOARES, S. S. D. *O Atrito Nas Pesquisas Longitudinais: O Caso da Pesquisa Mensal de Emprego (PME) do IBGE*. [S.l.], 2008. Citado na página 43.

SCHMUTTE, I. M. Job referral networks and the determination of earnings in local labor markets. *Journal of Labor Economics*, v. 33, n. 1, p. 1–32, 2015. Citado na página 18.

TOPA, G. Chapter 22 - labor markets and referrals. In: BENHABIB, A. B. J.; JACKSON, M. O. (Ed.). [S.l.]: North-Holland, 2011, (Handbook of Social Economics, v. 1). p. 1193 – 1221. Citado na página 15.

TOPA, G.; ZENOU, Y. Neighbourhood versus Network Effects. *Handbook of Regional and Urban Economics*, v. 5, p. 561–624, 2015. Citado na página 17.

Appendix

APPENDIX A – Tables

Table 15 – Nationality Frequency

Nation	Secondary Education		Nation	Superior Education	
	Freq	Percent		Freq	Percent
Bolivia	17,220	19.84	Portugal	11,004	12.68
Paraguay	4,594	5.29	Argentina	8,082	9.31
Haiti	2,973	3.43	Chile	7,582	8.74
Peru	1,331	1.53	China	4,574	5.27
Angola	321	0.37	USA	3,879	4.47
South African	189	0.22	Japan	3,614	4.16
Congo	86	0.1	Spain	3,529	4.07
Pakistan	43	0.05	France	3,384	3.9
			Italy	2,992	3.45
			Uruguay	2,567	2.96
			Germany	2,317	2.67
			South Korea	1,836	2.12
			UK	1,236	1.42
			Other European	782	0.9
			Switzerland	651	0.75
			Colombia	554	0.64
			Canada	411	0.47
			Belgium	376	0.43
			Venezuela	239	0.28
			Bangladesh	143	0.16
			India	115	0.13
			Equador	78	0.09
			Russia	75	0.09
Total	86,777	100			

RAIS dataset from 2009 to 2013. Percents are calculated from the whole foreigner sample.

Table 16 – Nationalities by Immigration Age Groups

First Quartile		Second Quartile		Third Quartile		Fourth Quartile	
Nation	Frequency	Nation	Frequency	Nation	Frequency	Nation	Frequency
Paraguay	4,594	Bolivia	17,226	Argentina	8,082	Portugal	10,999
Haiti	2,971	China	4,577	Chile	7,582	Japan	3,615
Colombia	554	USA	3,881	Uruguay	2,568	Spain	3,530
Equador	78	France	3,383	South Korea	1,834	Italy	2,992
		Peru	1,332	Switzerland	651	Germany	2,316
		UK	1,236	Angola	321	Other European	782
		Canada	411	India	115		
		Belgium	376	Russia	75		
		Venezuela	239				
		South African	189				
		Congo	86				
		Pakistan	43				

Note: Immigration Years are median values from the Census Survey from 2010.

Table 17 – Nationalities by Network Size Groups

First Quartile		Second Quartile		Third Quartile		Fourth Quartile	
Nation	Frequency	Nation	Frequency	Nation	Frequency	Nation	Frequency
Italy	2,992	Paraguay	4,594	Argentina	8,082	Bolivia	17,226
Haiti	2,971	China	4,577	Chile	7,582	Portugal	10,999
Uruguay	2,568	USA	3,881				
Germany	2,316	Japan	3,615				
South Korea	1,834	Spain	3,530				
Peru	1,332	France	3,383				
UK	1,236						
Other European	782						
Switzerland	651						
Colombia	554						
Canada	411						
Belgium	376						
Angola	321						
Venezuela	239						
South African	189						
India	115						
Congo	86						
Equador	78						
Russia	75						
Pakistan	43						

Note: Network size is defined by the frequency of each nation in the RAIS dataset.

Table 18 – Nationalities by region

Latin America	North America	Central Europe	Western Europe	Africa	East Asia	Other Asia
Bolivia	USA	Portugal	Other European	Angola	China	India
Argentina	Canada	Spain	Russia	South African	Japan	Pakistan
Chile		France		Congo	South Korea	
Paraguay		Italy				
Haiti		Germany				
Uruguay		UK				
Peru		Switzerland				
Colombia		Belgium				
Venezuela						
Equador						

APPENDIX B – PME Results

Since the PME has a direct variable of referral, we proposed two questions to be answered: do people that tried to find a job through referral had greater probability to succeed in the subsequent interview, and do people that got a job through referral stayed employed in the next two interviews more than those that found from another measure?

The data set that will be used is the PME (Monthly Employment Survey, in Portuguese). We will restrict our data to the Sao Paulo Metropolitan Area, or SPMA, in the years of 2007 to 2012. Because there is a lot of friction and loss between interviews, Ribas e Soares (2008) developed an algorithm to identify and reduce them, which we use to recover the longitudinal data.¹

In terms of regression analysis, the baseline equation is the following:

$$y_i = \alpha_i + \mathbf{d}\boldsymbol{\gamma} + \mathbf{X}\boldsymbol{\beta} \quad (\text{B.1})$$

Where y_i is whether individual i was employed or not, $\mathbf{d}$ is a vector of dummies which represent the measure the individual took to obtain the job and $\mathbf{X}$ is a vector of socioeconomic characteristics. The main variable of interest in this simple analysis is whether the individual took any measures to obtain a job in the last month, and which one was that.

To check for evidences of the two aforementioned questions, the data set has to be arranged in different ways. For the first one, whether people that contact friends and family have greater probability to find a job, we begin by keeping only those that were present and economically active throughout the first four interviews and were unemployed in the first interview while also looking for jobs in any of the above measures. In this way, if the agent answered in $t - 1$ it referred to relatives or friends, and in t got a job, we identify this as a referral hiring. Then, we maintain only the second interview, and regress the state of the individual (employed or not) against three dummy variables that are 1 according to the measure it took in $t - 1$ and some socioeconomic variables.

For the second question, the initial sampling is the same, but in this case, we restrict ourselves to those that were unemployed in the first interview, took a measure to find a job and got it in the second interview, or were employed in the first interview, took a measure to find a job and got it in the second interview by changing jobs. Then, we

¹ We are thankful to Data Zoom, developed by the Department of Economics at PUC-Rio, for providing the codes for accessing IBGE microdata.

check whether this individual stayed in this job for the next two interviews, and regress it against the dummies and socioeconomic variables.

Note that these models suffer from several identification problems, including low number of observations, and do not try to imply any kind of causality between the variables, but merely correlations and indications of the validity of the hypothesis. After subsetting for the two cases, the data set goes from 581,836 observations and 145,459 unique individuals to 854 and 100 observations respectively (also the number of unique individuals).

We would also like to see if there is a difference on the relevance of referral hiring between ethnicity and race, so we also regress using only White and Black individuals. Below are the results. Columns (1), (2) and (3) models 1 with full sample, only with whites and only with blacks respectively, while (4) is for model 2 with full sample. Since the number of observations dropped significantly for model 2, the separate regressions were not estimated.

Table 19 – Regression Results

VARIABLES	(1) Model 1	(2) White	(3) Black	(4) Model 2
Referral	0.124* (0.0714)	0.118 (0.0918)	0.0959 (0.115)	0.452* (0.247)
Agencies	0.0607 (0.0631)	0.153* (0.0901)	-0.0265 (0.0962)	0.433* (0.257)
Direct	0.0747 (0.0513)	0.0827 (0.0618)	0.0383 (0.0882)	0.300 (0.220)
Observations	854	451	399	100
R-squared	0.039	0.049	0.040	0.139

Robust standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Note: Included race, education, gender and age controls.

We see that, according to the results, the probability of getting a job is on average 12% higher than the other methods, which have no statistical significance. When we divide between whites and blacks, agencies seem to become the relevant method for the first group. When we look to the second model, we see that both referral and agencies seem to rise the probability of maintaining that job for the next two months.

Again, these indications are mere correlation. But, the above results might indicate referral hiring is relevant in the São Paulo labor market, and that there may be differences between races.