Informal Refferrals, Employment, and Wages: Seeking Causal Directions

A. M. Diaz E.

Discussion Paper 2009-23

Institut de Recherches Économiques et Sociales
de l'Université catholique de Louvain

UCL

INFORMAL REFERRALS, EMPLOYMENT, AND WAGES: SEEKING CAUSAL DIRECTIONS

ANA MARIA DIAZ E.[†]

Abstract. Employers and job seekers rely extensively on job informational networks to fill vacancies or to find a job. The widespread use of job contacts to find work has been largely associated with labor outcomes, such as finding a job or even affecting wages. Some scholars have claimed that informal referrals play a determinant role in reducing informational mismatches between potential employers and job seekers. Although several studies have shown that the use of friends and relatives is correlated with labor outcomes, little is known about the causal effect. In this article, I aim to identify whether there is a causal effect of using informal referrals on two main outcomes: the probability of being employed and hourly wages. I use a large data set from Colombia, the Living Standard Survey 2003, to contrast the results from three main methodologies: standard OLS estimation, propensity-score matching, and instrumental variables. Results suggest that much of the positive effect of using informal referrals on employment reflects the prevalence of informal-sector jobs to be filled through this method rather than a causal effect. On the contrary, the results for hourly wages suggest a negative causal effect of using job informational networks, which is explained by the low-quality/poor matches theory. Yet, this is only true in formal-sector firms.

Key words: informal referrals, job search methods, employment rates, hourly wages, selection bias, OLS, Propensity-Score Matching, Instrumental Variables, Roy Model.

JEL Classification: J21, J24, J31, J64, O17, C14, C14.

Date: August 11, 2009.
I am especially grateful to Bart Cockx for very helpful comments and guidance. This paper has also benefited from comments from Muriel Dejemeppe, Giordano Mion, Sofia Pessoa e Costa, and the seminar participants at the Doctoral Workshop, and the Social Capital conference 2008. Special thanks to Silvia Espinosa whose detailed suggestions made this text more readable and, hopefully, more useful. Any remaining errors, however, are my own responsibility. Finally, I am indebted to the International Relation Office (ADRI in French) of the University for financial support.
[†] IRES and Department of Economics, Université Catholique de Louvain, Place Montesquieu 3, 1348 Louvain-la-Neuve, Belgium. E-mail: `Ana.diazescobar@uclouvain.be`.

1. Introduction

Employers and job seekers are brought together through their recruitment and job search strategies. These strategies are classified as either formal or informal. The formal mechanism operates via media advertisements, and public or private intermediary agencies. Informal methods rely on information from personal intermediaries such as friends, relatives, or acquaintances.

The recent literature has investigated the relationship between job search strategies and labor outcomes. Specifically, a large amount of literature analyzes the effect of using informal referrals on labor outcomes.[1] Section 2 reviews this literature focusing on those studies that provide a theoretical framework of the mechanism through which the job search strategy affects both employment opportunities and wages; in this section I also survey the latest empirical studies. The main lesson from this literature is that informal search methods are widely used. Depending on the country, informal referrals are typically estimated to account for somewhere between 25 to 70 percent of hires (Calvó-Armengol & Jackson, 2004; Montgomery, 1991). The literature also acknowledges that friends, relatives, and acquaintances play a major role in solving information asymmetries and other frictions in the labor market and, hence, potentially affect labor outcomes. But, unfortunately, there is no consensus about the channel through which informal referrals operate to affect labor outcomes, particularly wages. Greater discrepancies are evident in the empirical literature. Indeed, sampling limitation, methodologies, and measurement problems make it difficult to draw firm conclusions about the extent to which there is a causal relationship between the use of informal referrals and labor outcomes.

In this article, I address several empirical questions concerning the *causal* link between job search choices and both the probability of being employed and wages, using the Living Standard Survey (LSS) for 2003, a national representative household survey in Colombia.[2] A preliminary data analysis of the LSS survey, Section 3, confirms that informal referrals are often used: around 40 percent of job seekers rely on this method to find a job, and roughly 70 percent of salaried workers report that they were hired through informal referrals. Moreover, a naive comparison of average outcomes shows significant differences between referred and non referred workers. There is a 21.4 percentage point difference between the employment rate of informal referral users (88.2

[1] For recent surveys refer to Ioannides & Loury (2004); Marsden & Gorman (2001)

[2] One should expect that agents seek for employment through informal channels whenever formal channels work poorly or are inexistent. Colombia seems to fit this description since there is not a centralized unemployment system, the only state employment agency performs badly and private employment agencies are limited.

percent) versus non-users (66.7 percent). Additionally, referred workers earn on average 44 percent less than non-referred workers. These comparisons are very likely to be biased if there exist both observed and unobserved attributes that jointly affect the decision of using informal channels and outcome variables. Such process is known as selectivity bias, which is a result of individual self-sorting, selection made by the employer, or both. For example, job seekers sort into job search strategies based on their own attributes (age, education, gender, and social abilities), which might be precisely the attributes that employers value in their selection and wage-setting processes. The measured referred/non-referred differentials, therefore, might simply reflect the effect of these attributes.

This paper contributes to the literature by attempting to identify causal effects of the use of informal referrals on both the probability of being employed and hourly wages, while explicitly taking into account the potential selectivity bias. Section 4 reviews the main empirical strategies that I use to circumvent this bias. I focus on three widely used econometric methods: ordinary least squares, propensity-score matching, and instrumental variables. The first two methods address the selectivity bias based on observable individual attributes. The last recognizes the presence of unobserved variables that affect both the use of informal referrals and the probability of being employed, even after conditioning on observed variables. It requires, however, the choice of an adequate instrument. I propose a variable that proxies for the use of informal referrals within the household, which is a binary variable equal to one if the closest blood relative in the household obtained his current job through informal referrals. This instrument conveys information about potential household correlation of the use of informal referrals, as well as potential channels of job related information. Thus, we can argue that it is highly correlated with the use of informal referrals and that its only correlation with both employment status and hourly wages is through the use of informal referrals. Yet, to avoid any remaining effect between the instrument and unobserved factors affecting the outcome it is necessary to control for observed attributes that might be correlated at the household level; in this case I control for employment attributes of the closest relative.

Section 5 presents and discusses the results of implementing these methodologies. In general terms, the results suggest that, after controlling for unobserved attributes, informal referrals are not more successful than any other job search strategy in terms of job seekers' placement. In other words, the employment rate differential between referred and non-referred workers is not significantly different from zero. I argue that much of the initial positive effect of using informal referrals is due to differences in the observed characteristics and selection effects.

On the contrary, the results suggest that informal referrals usage has a negative effect on hourly wages after controlling for all potential confounding variables (i.e., observed and unobserved variables). This finding is robust to minor variations of the control variables, and variations of the methodologies. Is important to clarify that I use the sub-sample of salaried workers to calculate the estimates, and the estimation procedures used in this section do not control for self-selection into employment. To explore whether this process biases the results I estimate the parameters from a two-stage Heckman model, and there is no evidence that the failure to control for self-selection yields biased results in this case. The point estimates suggest that referred workers earn on average 13 percent less than non referred workers.

This finding can be interpreted in at least two ways. One interpretation is that the use of informal referrals generates a mismatch between a worker's occupation choice and his comparative productive advantage because informal referrals induce both the average quality of the labor force and the return to firms' investment to remain low (Bentolila *et al.* , 2004). Thus, jobs found through contacts are obtained more quickly but also pay lower wages, since at least some of them are filled by workers who sacrifice their productive advantage in order to get a job more easily. An alternative interpretation is that informal referrals are proxying for unobserved job attributes. For example, jobs reachable through social networks are available only in firms that pay lower wages, regardless of skills. In Colombia, this is particularly the case of informal-sector firms where neither hiring nor minimum-wage laws are enforceable. Thus, the negative effect of informal referrals on wages might simply reflect that workers in the informal sector earn less than workers in the formal sector. If this is true and there is some aspect of informal sector jobs that is not adequately controlled for by observed variables, then the true effect of using job contacts may be biased downwards.

One way to try to get around this problem is to carry out a model known in the literature as the Roy model, which consists of two wage equations (one for each sector) and a selection equation that determines the sector in which the employee is working. The model supports the idea that referred workers have a higher probability of working in the informal sector, but within this sector there is no significant wage gap between referred and non referred workers. On the contrary, there is a wage discount of around 13 percent for jobs found through contacts in the formal sector. This result supports the theory that informal referrals tend to distort workers' occupational choices, inducing low quality matches. Yet, this is only true for formal-sector occupations.

2. Literature Review

In this section I use several recent review articles to divide the analysis into two main categories: the effect of informal referrals on employment, and their effect on wages. In each subsection, I first delineate the theoretical approaches and subsequently discuss the empirical findings.

Employment. Authors often link the use of informal referrals and labor market advantages by asserting that social networks reduce information asymmetries between job seekers and potential employers, and that they constitute a low-cost job search method (Montgomery, 1991; Mortensen & Vishwanath, 1994; Holzer, 1987b). Although job-related information can be acquired through formal methods, it can be obtained more rapidly through social networks (Aguilera, 2002). Moreover, in labor markets undermined by various sorts of frictions and long spells of unemployment, job seekers may use informal referrals in order to locate vacancies without bearing high search costs (Holzer, 1987a; Calvó-Armengol & Jackson, 2004; Fontaine, 2004). On the other hand, job seekers not using social networks may miss job opportunities only available through personal networks. For example, Montgomery (1991) shows that employers often delegate the screening function of finding the most suitable employees to the network of their current workforce because they are more likely to refer workers of the same type. Applicants outside the network do not learn about such jobs. Under these conjectures, one can argue that referred workers are in an advantageous position in the job finding process.

Recent studies have shown that the benefits from using informal referrals in the job search process cannot be generalized, since their usage might affect particular subgroups more than others (Calvó-Armengol & Zenou, 2005; Calvó-Armengol & Jackson, 2006; Ioannides & Loury, 2004). Their benefits might be shaped by institutional, employer, network, and worker heterogeneity. For instance, the institutional background might determine the conditions under which job search methods operate (Fontaine, 2004). Employer heterogeneity denotes the fact that some sectors, professions, and firms are more prone to rely on job informational networks to fill their vacancies (Corcoran *et al.* , 1980b; Holzer, 1987b; Marsden & Gorman, 2001). Network heterogeneity denote variations in the resource endowment of workers' contacts and also the relationship with the contact that might affect the transmission of job information (Calvó-Armengol & Jackson, 2004; Mortensen & Vishwanath, 1994; Lin, 2001; Ioannides & Loury, 2004).[3]

[3] The size of the network might play a relevant role since more social links lead to a more fluent transmission of information about employment opportunities (Granovetter, 1973; Calvó-Armengol, 2004). Aside from the size, the literature recognizes that the transmission of job information is affected by the quality of the network (Calvó-Armengol & Jackson, 2004; Mortensen & Vishwanath, 1994; Lin, 2001).

Finally, worker heterogeneity refers to differences in worker productivity or other relevant individual employability attributes. It interacts with all four of the other areas in determining access to contacts and employers (Ioannides & Loury, 2004).

Empirical analyses suggest that social networks do affect the eventual assignment of workers to jobs. Although it is not the particular interest of this paper, it is important to note that most of the empirical literature has evaluated the effect of informal referrals on unemployment duration. There is a general agreement in this literature that informal job search methods are associated with shorter unemployment duration and increased likelihood of exit when compared to formal job search methods (Montgomery, 1991; Blau & Robins, 1990; Korpi, 2001; Osberg, 1993). The few empirical papers examining the relationship between job search methods and employment pertain to specific groups, such as immigrants, blacks, women, or youths. Their results indicate an important variation in the relationship between informal referrals and employment across groups (Yakubovich, 2005; Bortnick & Ports, 1992; Alon & Stier, 2004; Chapple, 2002; Bian, 1997; Amuedo-Dorantes & Mundra, 2005; Borghans *et al.* , 2006; Goos & Salomons, 2007; Aguilera, 2002). For example, Bortnick & Ports (1992), using US data, find that men who were unemployed in a given month in 1991, and who used informal referrals that month, were slightly more likely than their female counterparts to find a job a month later. They also find racial differences, since whites were more likely to be employed a month later if they used help from friends and relatives than their black counterparts. Alon & Stier (2004); Chapple (2002); Borghans *et al.* (2006); Goos & Salomons (2007), using both different data and methods, confirm these results.

These empirical analyses are usually based on observational data and do not explicitly take into account the possible selectivity bias. One should be careful, therefore, to interpret these results as causal effects. For example, the fact that unrepresented groups (e.g., females, blacks, immigrants) are less likely to find a job when using informal referrals than their counterparts (e.g., males, whites, non migrants) might simply reflect that those groups are less likely to be employed regardless of which job search method they have used, and the estimates might simply be an artifact of the selectivity bias.

Wages. Diverse theoretical foundations provide varying (and sometimes conflicting) links between the job search method and the resulting wage. Evidently, empirical research trying to test such foundations also provides conflicting results. Here, I provide a review of relevant theoretical literature to be followed by a brief review of the main empirical findings.

Some researchers argue that the use of informal referrals in the job search process leads to higher wages. This literature states that although the use of informal referrals does not directly affect worker's productivity, it could improve the quality of the job match since both potential employers and job seekers have access to better and more reliable sources of information (Montgomery, 1991; Mortensen & Vishwanath, 1994; Simon & Warner, 1992). Employers can exploit information flowing within social networks to screen their potential applicants, and potential employees also have access to valuable information about non-pecuniary aspects of employment and, hence, might find jobs more closely matching their skills and preferences (Fernandez & Castilla, 2001). Another possible explanation for the referral wage premium is that employed contacts are more prone to refer only good applicants because their reputation is at stake and, therefore, less able workers will find it more difficult to get a reference in the first place (Kugler, 2003). For the same reason, high-effort referees can exert peer pressure on coworkers once they are hired, thus making them more productive. Yet, positive effects might also arise due to simple correlation induced by nepotism or favoritism instead of effects on workers' productivity. For example, Lin (2001) shows that employed contacts of unemployed workers may also directly influence the job-matching process by providing entry into highly paid jobs regardless of their qualifications. In the same way, Goldberg (1982) asserts that less competitive industries may be able to afford maximizing utility instead of profits and pay wage premiums to referred applicants.

In contrast to these views, some theories postulate that there is a negative relationship between the use of informal referrals and wages. This literature claims that job seekers are willing to sacrifice higher wages to obtain a position rapidly. For example, Bentolila *et al.* (2004) emphasize that job contacts may produce a mismatch between workers' comparative productive advantage and their occupational choices. Such conjecture implies that informal referrals create poor matches: the availability of informal referrals and the opportunity to find a job more easily may induce a job seeker to accept a job offer in professions, sectors, or locations where his abilities are not fully exploited. They show that people using social contacts have on average lower abilities and, therefore, would earn less than individuals who used other methods. However, the literature also argues that this negative effect can stem from unobserved attributes of either the worker or the job. For example, Granovetter (1995) argues that workers who rely on contacts are likely to be the ones who are in greater need of a job, and thus, they are likely to accept lower wages. So, the use of informal referrals can be a negative signal for employers who can respond by offering a lower wage. Additionally, firm attributes might also matter since jobs reachable through social networks are usually available in firms which pay lower wages (Labini, 2004). We can therefore summarize these two theories into causal and correlated effects. The causal effect argues that informal referrals result in a mismatch between a worker's productive comparative advantage and his occupational choices (induced by social ties), which in turn might affect the worker's productivity. On the other hand, the correlated effect argues that individuals who rely on informal referrals are more likely to earn lower wages because they are low ability individuals regardless of the job search method used, or because they are employed in firms that strongly rely on informal referrals to fill their vacancies but also pay lower wages.

Limited by the particular theory, methodology, and related variables chosen for analysis, the empirical literature often provides support to opposing theories. Empirical results from job information networks support a positive relationship between the use of informal referrals in the job search process and accepted wages. Granovetter (1973), Corcoran *et al.* (1980b), Simon & Warner (1992), Rosenbaum *et al.* (1999), Marmaros & Sacerdote (2002), Green *et al.* (1999), and Kugler (2003) find a positive wage premium for jobs obtained through referrals from current employees. However, empirical support for significant positive effects is not universal. Elliot (1999), Datcher Loury (2006), Bentolila *et al.* (2004), and Pellizzari (2004) report a negative relationship between social networks and wages. Finally, Bridges & Villemez (1986) and Mouw (2003) claim that the use of friends and relatives in the job search process does not have any effect on wages.

The absence of experimental data makes us cautious about interpreting these results when it comes to distinguishing between causality and correlation. For example, a positive wage gap between referred workers and non referred workers might result from good-quality matches which improve worker productivity. On the contrary, a negative wage gap between referred and non referred workers might result in low/poor-quality matches which will deteriorate worker productivity. These two cases can be understood as causal effects. It turns out, however, that those who do not use informal referrals are also more likely to have specific individual attributes that directly affect their performance and, therefore, their wages; it is the worker's skills that are playing a significant role in their performance and wages. Moreover, the informal referrals variable might be proxying for unobserved job attributes. For example, jobs where informal networks work effectively (in the sense of successfully matching job seekers to jobs) may be available only at firms that pay lower wages on average irrespectively of their workers' abilities. Concerns over such sources of bias on estimates of returns to education have resulted in a large literature in economics that attempts to identify the causal effect of education on wages and income (see the extensive review in Card (1999) and Blundell *et al.* (2005)). However, the empirical literature on social networks has failed to control for the presence of this bias. One exception is the paper of Mouw (2003), where he replicates several empirical analyses for the US using a fixed-effects model to difference out fixed unobserved individual attributes, and he concludes that there is no evidence to argue that using contacts causally affects wages.

3. Data Description

The empirical analysis is based on Colombian data from the Living Standards Survey (LSS) of 2003, a nationally representative household survey carried out on a sample of 22,090 households, from which 82,495 individuals were interviewed (see Appendix A for a more detailed description of the survey)[4]. The particular sample used in this analysis is restricted to the labor force living in urban areas, which corresponds to 22% of the aggregate sample. It also excludes self employed and non-paid workers since they do not report their job search strategies.

Table 1 presents some relevant figures of these data. Around 80 percent of the labor force is employed, and earns on average almost one US dollar per hour. The Table also reports descriptive statistics for the labor force, and separately for salaried workers, and depicts the distribution of each sample by relevant observed characteristics used in the empirical analysis. On the basis of the figures, one can deduce that the composition of

Table 1. Descriptive Statistics, LSS 2003.

Attribute	Active Population Excludes SE* Proportion	Salaried Workers Proportion
Main Outcomes		
Employed	80.02%	
Hourly Wage (log)		US\$0.95**
Individual Attributes		
Gender		
Male	52.26%	55.30%
Female	47.74%	44.70%
Level of Education		
Primary or Less	14.89%	14.82%
Secondary	44.24%	43.15%
Higher Educ non Univ.	11.29%	11.83%
Uncomplete University	10.40%	10.11%
Complete University	11.79 %	12.57%
Post University	7.18%	7.53%
Age Range		
15 - 24	21.11%	19.14%
25 - 34	31.39%	34.09%
35 - 44	24.64%	27.48%
45 - 54	14.48%	15.37%
55 - 64	4.37%	3.92%
Ethnics		
White	94.05%	94.36%
Single	39.34%	34.67%
Household Attributes		
Unipersonal Household	4.65%	4.80%
Presence Child 0 - 4	14.34%	15.77%
Presence Child 4 - 6	13.51%	14.78%
Presence Child 7 - 12	25.83%	27.02%
Closest relative employed	57.11%	55.82%
Instrument		
Closest relative employed by IR	21.08%	21.19%
Observations	18.049	11.468

NOTE: Standard Deviation in Brackets, Self-employed also includes non paid workers. * SE accounts for Self employed workers. ** Mean of Hourly wages corresponds to 2,643 pesos which represents US\$0.95 (exchange rate 2.778 pesos/US dollar in 2003)

[4] The figures and results depicted in this section were contrasted with the Continuum Household Survey for 2002, 2003, and 2004. The results, available upon request, are similar to those reported in this section.

the population by both individual and household attributes does not vary across samples. The only relevant distinction is that the salaried sample has a larger fraction of males and middle aged individuals than the active population sample. In sum, the samples are mainly composed of males, with secondary education, who are middle aged, white and non-single.

The Treatment Variable. The LSS collects information regarding the main job search method for both unemployed and salaried workers. The survey asks unemployed workers: "What was your main activity in the last four weeks to find a job?". For salaried workers, the question is: "How did you find your current job?". Both salaried and unemployed can select among one of the following eight channels: i. ask for help from relatives, friends or ex-colleagues, ii. from the governmental center for employment information, iii. direct application, iv. visit private agencies, v. by publishing adverts in newspapers, vi. public announcements, vii. Internet, and xiii. tried to establish your own business.

Table 2 compares the response distribution for the unemployed and the employed groups. Although both of them rely on informal contacts to search for a job, the proportions vary considerably across groups. Seventy percent of salaried workers report that they were hired using help from contacts, while only forty percent of unemployed focus their job search process on this method.

Table 2. Job Search Methods, Percentage by Employment Status, LSS 2003.

Method	Unemployed	Salaried
Informal Referrals	39.32%	69.30%
Direct Applications	35.35%	13.81%
Government	2.22%	0.90%
Private Agencies	6.82%	3.33%
Advertising	1.24%	2.66%
Public	1.05%	5.90%
Internet	1.00%	0.20%
Other	13.01%	3.89%

Source: Own's calculations from LSS 2003

Using these data I define a binary variable that describes the use of informal referrals, which I name IR. It takes the value of one in two situations: i. if the individual is unemployed and declares informal referrals as his main strategy to find a job, or ii. if the person is employed and states that he has found his current job through informal referrals. The IR variable takes the value of zero if the individual is either unemployed or employed and declares the use of other channels to find his job.

A major disadvantage of the LSS survey is the lack of information on whether the individual is seeking for a job through different channels simultaneously. Individuals do not necessarily focus their search on a single channel, but rather are more likely to diversify their search in order to increase the probability of obtaining a job. Ideally, one would prefer to analyze both the use of informal referrals as a single job search method and mixed strategies of using a combination of job-search methods. Lacking this information, I restrict the analysis to the main channel.[5]

To conclude this section, I describe the observed difference in both outcomes by job search channel. Table 3 reports the average for both employment and hourly wage by job search method, as well as, the *naive* difference between them. It can be seen that informal referral users have a higher employment rate than non users (a difference of 21 percentage points). On the contrary, referred workers earn on average 44 percent less than non referred workers.

Table 3. Labor Market Outcomes by Job Search Method, LSS 2003.

Outcome	IR user	IR non user	Difference
	(1)	(2)	(1-2)
Employment Rate	88.15%	66.69%	21.46pp
S.E.	[0.30%]	[0.56%]	[0.59]
Hourly Wages (log)	7.74	8.18	43.78%
S.E.	[.013]	[.008]	[0.015]

Source: Own's calculations from LSS 2003

4. Empirical Methodology

This section describes the methodologies used in the empirical analysis. It uses the counterfactual approach widely used in the microeconometric literature (Heckman *et al.* , 1997; Heckman, 1997; Heckman *et al.* , 1998; Imbens & Angrist, 1994) to initially describe the fundamental estimation problem, and to subsequently discuss the identifying assumptions of each econometric strategy. It then briefly delineates how these strategies are implemented to yield practical estimators.

To illustrate the problem, let informal referral usage be described by a binary variable $IR_i = \{0, 1\}$. For notational simplicity, let us assume that both outcomes of interest (employment status and hourly wages) are denoted by y_i. Hence, for each individual there are two potential outcomes: i. the outcome if he uses IR, y_{i1}, and ii. the outcome if he uses other methods, y_{i0}. We would like to know the difference between y_{i1} and y_{i0}, which can be said to be a causal effect of using informal referrals for each individual, i.e., $\gamma_i = y_{i1} - y_{i0}$. However, only one of the potential outcomes is observed for each individual. Consequently, estimating the individual treatment effect γ_i is not feasible and one has to concentrate on average treatment effects by comparing IR users to a similar group of individuals who did not use this channel to obtain their job. Here, I focus on the widely used average treatment effect on the sub-population that uses help from friends and relatives, known as the average treatment effect on the treated (ATT), i.e., $E[\gamma_i | IR_i = 1]$.

[5] Another minor disadvantage is that the LSS does not include information regarding individual attributes of the contact. I cannot, therefore, evaluate whether the "quality" or social attributes of the contact affect labor outcomes. I instead argue that the primary advantage of contacts is to provide information about job openings, and such information will be useful despite the attributes of the person who provides it.

A naive comparison of observed outcomes between IR users and non IR users, shown in Section 3, provides us with the ATT plus a bias term:

$$\underbrace{E[y_i|IR_i=1]-E[y_i|IR_i=0]}_{\text{Outcome's Observed Difference}}=\underbrace{E[y_{i1}-y_{i0}|IR_i=1]}_{\gamma^{ATT}}+\underbrace{E[y_{i0}|IR_i=1]-E[y_{i0}|IR_1=0]}_{\text{Selection Bias}}$$

Indeed, the selection bias arises because the counterfactual, $E[y_{0i}|IR_i = 1]$, is not observable and any inference from observational data might yield biased estimates. For example, if individuals with favorable labor market attributes were more likely to have chosen informal referrals as their main channel to seek employment, it is likely that IR users would have done better on average than non IR users, regardless of which job search method they used. If so, the last term of the right hand side is positive and a naive comparison exaggerates the benefits of using informal referrals.

Many schemes have been proposed to circumvent this bias. Here I focus on three widely employed methods: standard regression methods, propensity-score matching, and instrumental variables. Although these methodologies try to address the plausible bias, each approach invokes different identifying assumptions to construct the required counterfactual outcome. I begin by considering both ordinary least squares and propensity score matching, which are based on the selection on observables assumption. Following that, I describe the instrumental variable methodology that allows for selection on unobservables.

Ordinary Least Squares (OLS) identifies the causal effects, γ^{ATT}, under the assumption that after controlling for observable characteristics, X_i, IR and the potential employment status are independent. This assumption is known in the literature as the conditional independence assumption (CIA), and it implies that if one can control for observable differences in individual attributes between IR users and non IR users, the outcome that would result without treatment is the same in both cases. In other words, the counterfactual equals the observed outcome for IR users, conditional on X_i, i.e., $E[y_{i0}|X_i, IR_i = 1] = E[y_{i0}|X_i, IR_i = 0]$.[6] Given this condition, the ATT can be constructed by iterating expectations over X_i,

$$\begin{aligned}\gamma^{OLS} &= E\left[E[y_i|X_i, IR_i=1]-E[y_i|X_i, IR_i=0]\right]\\ &= \underbrace{E\left[E[y_{i1}-y_{i0}|X_i, IR_i=1]|IR_i=1\right]}_{\gamma^{ATT}_{(x)}}\end{aligned}$$

[6] Rosenbaum & Rubin (1983) and numerous predecessors describe this assumption as $y_{i1}, y_{i0} \perp IR_i|X_i$. Heckman *et al.* (1997) argue that for identification of ATT a weaker condition is required, namely $y_{0i} \perp IR|X$. The only assumption that is required is the conditional mean independence assumption: $E[y_{i0}|X_i, IR_i = 1] = E[y_{i0}|X_i, IR_i = 0] = E[y_{i0}|X_i]$. It implies that y_{i0} does not determine the use of informal referrals.

OLS identifies the average treatment effect on the treated (ATT), under the assumption that the observed attributes, X_i, are the only reason why potential outcomes differ by job search channel. Therefore, controlling for them makes the selection bias disappear.

Propensity Score Matching (PSM) also assumes that the selection bias can be expressed purely in terms of observable characteristics. Similar to OLS, it uses observed explanatory variables to adjust for differences in the employment status unrelated to the use of informal referrals that give rise to selection bias. Therefore, for identification of causal effects, this method also relies on the conditional independence assumption (CIA).

Matching consists of finding a set of non-treated observations with the same realization of X_i for each IR user. A practical implementation problem arises when the vector X_i is highly dimensional. To circumvent this difficulty, Rosenbaum & Rubin (1983) show that matching on a scalar function of X_i, such as the propensity score, $P(X) = Pr(IR_i = 1|X_i))$,[7] is sufficient to balance the covariates X_i between the IR users and control units. Thus, if CIA holds conditional on X_i, it will also hold conditional on the propensity score, i.e., $(y_{i1}, y_{i0}) \perp IR_i|P(X_i)$. In order to ensure that CIA has empirical content, matching requires the existence of users and non users for each variable we seek to compare, i.e., $0 < P(IR_i = 1|X_i) < 1$. This assumption is known in the literature as the common support assumption.[8] Under these two assumptions, the ATT [9] can be expressed as,

$$
\begin{aligned}
\gamma^{PSM} &= E\left[E[y_{i1}|P(X_i), IR_i = 1] - E[y_{i0}|P(X_i), IR_i = 0]|IR_i = 1\right] \\
&= E[\gamma^{ATT}_{P(X)}|IR_i = 1]
\end{aligned}
$$

Although OLS and PSM rely on the CIA assumption, they differ in two main aspects. First, matching, by construction, eliminates the bias from having different ranges of X_i for the samples of IR users and non IR users (comparing non-comparable individuals, failure of the common support assumption) and the bias resulting from having different distributions of X_i across their common support (Heckman *et al.* , 1999). Second, matching is non-parametric, thus it avoids the restrictions involved in models that require the relationship between characteristics and outcomes to be specified.

7 The conditional probability of using informal referrals given the vector of observed covariates.

8 Failure to satisfy this assumption restricts the analysis to the region of support (all possible values of X_i) common to all IR users and non IR users, and the estimated treatment effect has to be redefined as the mean treatment effect for those treated falling within the common support region.

9 The empirical counterpart of this equation is:$\gamma^{\hat{ATT}}_{PSM} = \frac{1}{N_1}\sum_{i \in IR=1}\left\{y_{1i} - \sum_{i \in IR=0} W_{N_0} y_{0i}\right\}$. N_1 is the number of IR users and N_0 is the number of non IR users, W_{N_0} is a weight function that weights the observations of an IR non-user according to their similarity with the observed covariates of the IR user. In this case I use Kernel Matching, where IR users are matched with a weighted average of all controls using weights that are inversely proportional to the distance between the propensity scores of IR users and non IR users (Becker & Ichino, 2002; Todd, 2006).

PSM and OLS estimates assume that, conditional on the individual attributes that employers use to select their employees and to set wages, the use of informal referrals is independent of potential outcomes. The motivation to use these strategies is the assumption that employers screen applicants and set wages primarily on the basis of observable covariates like age, schooling, gender, and race. Thus, after controlling for all these attributes, IR users and non-users are comparable. This assumption holds, however, if and only if there is no omitted variable bias once X_i is included in the regression; in other words, it will hold if there are no unobserved characteristics that influence the use of informal referrals which are also associated with the outcomes. This assumption, however, cannot be formally tested. Instead, I will compare the estimated results from both methodologies with those obtained by instrumental variables.

Instrumental Variables recognizes that workers who use informal referrals are likely to be non-randomly selected, even conditioning on observed attributes. Heckman (1997) describes this method as a variant of the method of matching, since it augments the X variables in matching with instruments Z. Such instruments must satisfy the following conditions: i. the use of informal referrals depends in a non-trivial way on both X_i and Z (i.e., $E[IR_i|X_{iy}, Z_i] \neq E[IR_i]$), and ii. the instrument does not have any direct impact on individual employment status, conditional on the observed individual attributes (i.e., $E[y_{i0}|X_i, Z_i] = E[y_{i0}|X]$).

Here, the variable used to instrument for the use of informal referrals is a proxy of their usage within the household. It is a binary variable, $Z_i = \{0, 1\}$, that indicates whether the closest blood relative within the household obtained his current job using informal referrals. From Boorman (1975) and Calvó-Armengol (2004) we can assume that the way in which the job related information flows within the network is the following: if an employed individual hears about a job opening, and it is not interesting for himself, he passes it on to his closest relative. Therefore, the instrument conveys information about potential household correlation on the use of informal referrals, as well as on potential channels of job related information. However, it does not contain information about individual-specific chances to find a job resulting from unobservable characteristics such as social skills, and should therefore be uncorrelated with unobserved employment components. Under this assumption one might justify that the instrument can only affect the employment situation through the use of informal referrals.

A skeptical reader can find reason to doubt the validity of this instrument since Z itself is endogenous and confounded with IR and y_i. The use of informal referrals at the household level might induce some individuals to also use this method, but it is unlikely to affect both employment and wages directly. Nevertheless, the instrument is not randomly assigned but chosen by the closest relative in the household. Their choice, however, might itself be related to characteristics that directly affect their closest relative's subsequent employment status and wages. It is therefore necessary to control for these confounding covariates, X_{CR}, to handle the endogeneity of the instrumental variable Z. In this particular case, I include employment attributes of the closest relative such as firm size, position, sector, and whether he is a white collar worker.

In this framework, the engine that drives causal inference is the instrument $Z_i = \{0, 1\}$, but the variable of interest is still IR_i. Thus, the dependence of the employment status on the instrument arises through the dependence of IR on Z_i as follows,

$$E[y_i|Z_i = z] = E[y_{i0}|Z = z] + E[y_{i1} - y_{i0}]E[IR_i|Z = z]$$

Given that there are two possible realizations of Z, we can condition not only on the same set of observed attributes employed in OLS and PSM, X_y, but also on the employment attributes of the closest relative, X_{CR}, to distinguish two expectations $(E(y_i|X_{yi}, X_{CR}, Z = 1)$ and $E(y_i|X_{yi}, X_{CR}, Z = 0))$, and obtain the following IV (or Wald) estimator,

$$\gamma^{2SLS} = \frac{E[y_i|X_{yi}, X_{CR}, Z_i = 1] - E[y_i|X_{yi}, X_{CR}, Z_i = 0]}{E[IR_i = 1|X_{yi}, X_{CR}, Z_i = 1] - E[IR_i = 1|X_{yi}, X_{CR}, Z_i = 0]}$$

Under the assumption that Z_i has a clear effect on IR_i and that the only reason for the relationship between the employment status and the instrument is the first stage (once we control for the employment attributes of the closest relative), 2SLS identifies the causal effect of the use of informal referrals on both labor outcomes, whenever the selection is on unobservables. [10] [11]

In sum, each methodology aims to identify causal effects while explicitly taking into account the potential selectivity bias that might arise from individual self-sorting, selection made by the employer, or both. OLS and PSM assume that there are no important variables apart from X_i on which we cannot condition, that affect both the non-users employment status, y_{i0}, and the use of informal referrals, IR_i. If this assumption does not hold, the selection would be on unobservables and IV is called for, where identification of causal effects is achieved through Z conditional on the employment attributes of the closest relative.

[10] In the case of employment I obtain consistent estimates by carrying out OLS in both stages. As Angrist (2001) argues, conventional 2SLS estimates using a linear probability model are consistent whether or not the first stage conditional expectation function is linear.

[11] As Imbens & Angrist (1994) have shown, the standard interpretation of this estimator applies only when the treatment effect is constant among the treated. In the more realistic case of heterogeneous causal effects, and under a set of additional assumptions (i. stable unit treatment values (SUTVA), ii. ignorable assignment to treatment, and iii. monotonicity.), IV estimates the average treatment effect among those who modify the job search channel because the closest member did it; they call this parameter the local average treatment effect (LATE).

5. Empirical Results

This section presents some evidence on the empirical effect of informal referrals on both labor outcomes: the probability of being employed and hourly wages. Given that each outcome has its owns specificities, I analyze them separately. I explore the sensitivity of the point estimates to the exclusion of the closest relative's employment attributes, X_{CR}, by comparing two main specifications. I also examine whether the use of informal referrals affects particular subgroups more than others.

Employment. The first relevant outcome is the employment status of the active population, a binary variable equal to one if individual i is employed in the reference week. The treatment variable, IR, represents contact usage as the main method to seek for employment. Initially, in Specification 1 I control for the set of variables that directly affect both the probability of being employed and the labor supply decisions. This specification includes individual-level variables, household-level variables, the employment status of the closest relative, and regional dummies to capture any unobservable region-specific effects.[12] To explore whether the results are robust to the inclusion of additional family attributes that might be correlated within the household, in Specification 2 I add an additional set of employment attributes of the closest relative such as position, sector, firm size, and a white collar dummy.

The results, depicted in Table 4, suggest that there might a be confounding influence of the unobserved variables after conditioning on all observed attributes. This implies the violation of the CIA and, therefore, IV is called for. In fact, when this method is used, the predicted effect of informal referral is considerably lower than that predicted by OLS and PSM (see the last column). Moreover, the standard errors tend to increase, resulting in the failure to reject the null that informal referrals have a causal effect on the probability of being employed.

The OLS estimates indicate that referred individuals are about 23 percentage points more likely to obtain employment than non referred workers, and the coefficients do not vary across specifications. This result is supported by propensity-score matching (PSM), since the average treatment effect on the treated is very close to that obtained by OLS, with the only difference that they are less precisely estimated since they present larger standard errors. Like OLS, the estimates using PSM do not vary across specifications,

[12] The individual-specific covariates in this regression consist of dummies for male, marital status, race, age groups, and level of education. The household-specific covariates include dummies for the presence of children in the household, interactions of those with a male dummy, the presence of individuals older than 65, and a dummy variable that represents whether the individual lives alone. The employment status of the closest relative is comprised of two binary variables that describe whether he is employed or inactive.

suggesting that the use of informal referrals is not correlated with the employment attributes of the closest relative.[13]

Table 4: Effects of Informal Referrals on Employment, LSS 2003.

	OLS		PSM		2SLS	
	Spec 1	Spec 2	Spec 1	Spec 2	Spec 1	Spec 2
	$\hat{\gamma}_{X_1}$ [R-SE]	$\hat{\gamma}_{X_2}$ [R-SE]	$\hat{\gamma}_{P(X_1)}$ [BS-SE]	$\hat{\gamma}_{X_2}$ [BS-SE]	$\hat{\gamma}_Z$ [R-SE]	$\hat{\gamma}_Z$ [R-SE]
γ_0	0.232	0.229	0.235	0.233	0.26	0.051
	[0.006]***	[0.006]***	[0.007]***	[0.007]***	[0.082]***	[0.083]
Constant	0.412	0.418	0.647^a	0.649^a	0.393	0.422
	[0.020]***	[0.020]***			[0.057]***	[0.058]***
Control Variables	X_1	X_2	X_1	X_2	X_1	X_2
R^2	0.149	0.152			0.147	0.106
N	18049	18049	11189^b	11194^b	18049	18049
$Matched^c$			99.99%	99.99%		
Hausman -test^d					0.123	4.938
Hausman -test p-value					0.726	0.026
Weak Ident Test^e					91.34	89.655

* p<.10, ** p<.05, *** p<.01
X_1 = [Male, White, Single, Age, Educ, Region, Household Variables, CR employment status]
X_2 = [Male, White, Single, Age, Educ, Region, Household Variables, CR employment status, CR employment attributes]
NOTE: The table reports coefficients of probability of employment. Robust standard errors are in parenthesis. For each specification the first column shows the results from OLS, the second from propensity-score matching, and the third from two stage least squares. For results including the entire set of covariates refer to Appendix D Table 9. a. 722 individuals excluded from the analysis for non response, b. Treated Individuals, c. Percent of treated with at least one identical match in the control group, d. the endogeneity test is a C or distance GMM test, it produces a Hausman test that is robust to violations of homoskedasticity, e. Weak identification test reports the Wald F test robust to violations of homoskedasticity.

A simple way to examine the extent to which "selection on unobservables" may bias these results is to compare them with those from IV (last two columns).[14] The first specification does not include measures of employment attributes of the closest relative other than his employment status, and despite the fact that the effect seems to be slightly larger than the OLS estimates, we cannot reject the hypothesis that both coefficients differ by using a Hausman test robust to violations of homoskedasticity.

As mentioned in Section 4, it is very likely that the exclusion of employment attributes of the closest relative might be biasing this result since plausible correlation between the instrument and the individual employment status might arise. Indeed, when this information is accounted for in the second specification, the predicted effect of IR becomes considerably lower and less precise. Under this specification we can reject the hypothesis that the OLS and IV coefficients differ, moreover, the instrument is not

[13] Results of the imbalance covariate test (excluded here) show that the sample differences in the original data significantly exceed those in the sample of matched cases. Therefore, PSM creates a high degree of covariance between treatment and control samples. Additionally, the are no problems related with the common support assumption since almost all observations were matched.

[14] Recall that the variable used to instrument for IR is the same across specifications: a binary variable equal to one if the closest relative in the household found his current job by informal referrals. The use of informal referrals differs significantly if the closest blood relative also used this method: out of those whose closest relative used IR to obtain his current job, 71.51 percent declares the use of IR, as opposed to 59.16 percent of those whose closest relative used formal channels to obtain his current job.

weak.[15] Notice that the exclusion restriction only holds when we control for household attributes; hence, the results support that using informal referrals as the main method to seek for a job does not improve the probability of being employed compared to any other job search strategy.

There might be, however, differential effects by relevant observed attributes and, therefore, I test whether the use of informal referrals affects particular subgroups more than others. I use the methodology described in Wooldridge (2003), which amounts to allowing the IR effect to vary according to relevant individual attributes (i.e. gender, race, age, and education) while controlling for selectivity bias; for methodological details refer to Appendix B.

Table 5: Effects of Informal Referrals on Employment, Interaction Effects, LSS 2003

Var (X)	Specification 2 beta [R-SE] (1)	 Diff (X-Ref) (2)
ATT	**0.066** [0.123]	
Gender		
Female (ref)	0.207 [0.221]	
Male	0.387 [0.212]*	0.180 [0.051]***
Age Groups		
15-24 (ref)	0.207 [0.220]	
25-34	0.036 [0.208]	-0.172 [0.089]*
35-44	-0.037 [0.210]	-0.245 [0.091]***
45-54	-0.114 [0.204]	-0.321 [0.094]***
55-64	-0.259 [0.219]	-0.467 [0.123]***
Level of Education		
Primary or less (ref)	0.207 [0.220]	
Secondary	0.237 [0.164]	0.030 [0.161]
Higher educ non university	-0.103 [0.188]	-0.310 [0.202]
Unfinished University	-0.115 [0.195]	-0.322 [0.214]
University	-0.152 [0.172]	-0.359 [0.188]*
Post-University	-0.469 [0.179]***	-0.676 [0.201]***
Constant	0.422 [0.153]***	
R^2	0.08	
No of Observations	18049	

* p<.10, ** p<.05, *** p<.01

NOTE: The table reports coefficients on the probability of employment. Robust standard errors are in parenthesis. Column (1) reports the estimated effect of each category evaluated at the reference value of all the other variables. Column (2) reports the difference between each category and the reference category. See Appendix D Table 10 for the entire set of covariates

The results, depicted in Table 5, show significant differences between employment rates of referred and non referred workers according to gender, age, and level of education. Referred males have a higher probability of being employed than their female counterparts. Age related differences are also noticeable, since referred workers aged between 15 and 24 have a higher probability of being employed when using informal referrals than any other age group. Finally, referred workers with either university of postgraduate studies have a lower probability of being employed than individuals holding at most a primary school degree.

A possible reason for these results is the tendency for vacancies in the informal sector to be filled by informal referrals. The ideal scenario to test this assertion requires the

[15] Notice that the proposed instrument affects a specific population, i.e., those individuals whose closest relative is employed. Thus, to avoid extrapolation I carry out the same analysis for this sub-sample. Results for both OLS and PSM tend to be slightly larger since the effect accounts for 26 percentage points, which might reflect a selection effect. The results from 2SLS confirm the predictions described here: the inclusion of employment attributes of the closest member lowers the point estimates and it is no longer significant. These results are available upon request.

knowledge of the sector in which unemployed workers are seeking for jobs. Lacking this information, we can focus on employed workers to explore the data. Indeed, the figures from this sample confirm that proneness: 88 percent of informal-sector workers declare they have obtained their job through informal referrals, from which 63 percent are males, 42 percent are young individuals, and 28 percent are low educated (compared to 3 percent that has higher education).

Numerous papers have analyzed the empirical effect of using informal referrals on labor outcomes and have concluded that referred individuals are more likely to find a job than their non referred counterparts. Instead of showing a causal effect, this result might only reflect a correlation that arises from the high reliance on this method by both potential employers and employees. Indeed, using specific data and methodologies, my results suggest that the difference between employment rates for referred and non referred workers might simply be an artifact of selection bias, since when I control for unobserved confounding variables the effect vanishes. Moreover, the results also indicate the presence of heterogeneous effects, which are very likely to arise from employer characteristics that determine the context in which job search methods operate. Specifically, for a country like Colombia where the labor market is segmented into formal and informal sectors, and the later is more likely to rely on informal referrals, the fact that male, young, and low educated individuals present a higher probability of being employed if referred reflects that they are more likely to be hired in the informal sector and not by the usage of friends and relatives to seek for a job.

Wages. The second outcome of interest is the hourly wage, which is the monthly remuneration in the primary employment divided by the number of working hours. The "treatment" variable, IR_i, differs from that used in the employment case in the sense that now it is equal to one only for those who obtained their job through IR. To explore the robustness of the point estimates, I compare the results from two main specifications that differ on the set of included covariates. Specification 1 includes individual-specific covariates: male dummy, race dummy, age group dummies, attained educational level, and regional dummies. It also includes a binary variable representing whether the individual lives alone, and another one representing the employment status of the closest relative in the household. As for employment, Specification 2 augments Specification 1 to include an additional set of employment attributes of the closest relative.

Unlike employment equation estimates, the results from the wage equation do not suggest that, conditional on the observed covariates, the process by which workers use informal referrals is related to unmeasured variables that affect hourly wages. This conclusion arises from the impossibility of rejecting a difference between the OLS and the 2SLS estimators. The main implication of this result is that employers set wages only on the basis of productive individual attributes like schooling, age, gender, and race. Thus, after controlling for all these observed covariates, there are no unobserved individual attributes, such as communicative skills, that might modify the wage setting between referred and non-referred workers.

Table 6: Effects of Informal Referrals on Hourly Wages (log), LSS 2003.

	OLS		PSM		2SLS	
	Spec 1	Spec 2	Spec 1	Spec 2	Spec 1	Spec 2
	$\hat{\gamma}_{X_1}$ [CL-SE]	$\hat{\gamma}_{X_2}$ [CL-SE]	$\hat{\gamma}_{P(X_1)}$ [BS-SE]	$\hat{\gamma}_{(P(X_2)}$ [BS-SE]	$\hat{\gamma}_Z$ [CL-SE]	$\hat{\gamma}_Z$ [CL-SE]
γ_0	-0.134	-0.131	-0.155	-0.153	-0.138	-0.124
	[0.012]***	[0.012]***	[0.016]***	[0.017]***	[0.156]	[0.098]
Constant	7.304	7.271	7.847^a	7.693^a	7.307	7.265
	[0.033]***	[0.034]***			[0.116]***	[0.076]***
Control Variables	X_1	X_2	X_1	X_2	X_1	X_2
R^2	0.511	0.519			0.511	0.519
N	11468	11468	7771^b	7758^b	11468	11468
$Matched^c$			100%	99.98%		
Hausman Test^d					0.001	0.005
Hausman Test p-value					0.977	0.943
Weak Ident. Test^e					53.73	85.47

* p<.10, ** p<.05, *** p<.01
X_1 = [Male, White, Age, Education, Region, CR employment status]
X_2 = [Male, White, Age, Education, Region, CR employment status, CR employment attributes]
NOTE: The table reports the coefficients of hourly wages in logarithms. Clustered standard errors are in parenthesis for OLS and IV, the clustering in this case was done by household. Bootstrapped standard errors for propensity-score matching. The first column shows the results from OLS, the second from propensity-score matching, and the third from two stage least squares. For results including the entire set of covariates refer to Appendix D Table 11. a. Upper and lower 1 percent of the hourly wage distribution were excluded to avoid outliers, b. treated individuals, c. percent of treated with at least one identical match in the control group, d. the endogeneity test is a C or distance GMM test, it is robust to violations of homoskedasticity, e. Weak identification test reports the Wald F test robust to violations of homoskedasticity

Indeed, Table 6 shows that including measures of "productive" individual attributes considerably affects the observed wage gap between referred and non referred workers; recall from Section 3 that it accounts for about 43 percent less. Yet, when all the determinants of hourly wages are accounted for, the estimation becomes less negative. OLS results show that referred individuals earn 13 percent lower hourly wages on average, and this result is robust across specifications. Additionally, the point estimates from the propensity-score method confirm this conclusion.[16] Note however that the estimates of the ATT through PSM tend to be slightly more negative than those obtained through OLS, which might suggest the presence of heterogeneous effects.[17] The last two columns

[16] The results of the imbalance covariate test (excluded here) show that the sample differences in the original data significantly exceed those in the sample of matched cases, which creates a high degree of covariance between treatment and control samples. Additionally, the common support requirement was enforced, and the empirical estimation of the ATT ensured the existence of potential matches in the non users group for almost all IR users.

[17] Table 12 (Appendix D) shows the interaction effects between the use of informal referrals and main individual attributes. There are no differences by gender nor race. There are, however, differences by age and education. By age, it is evident that referred workers aged between 54 and 65 earn on average 10 percent less than those aged between 15 and 25. In the case of education we observe significant differences for low educated referred workers, particularly for referred workers with secondary education, and those that have some higher education but do not have a diploma.

show the 2SLS point estimates; as mentioned before, they do not differ from the OLS estimates even when controlling for the employment attributes of the closest relative. Moreover, the 2SLS estimates are less precise than those from OLS and PSM, as the standard errors rise considerably.

These estimates were carried out from a sub-sample of working individuals without any controls for self-selection into employment. This sub-sample contains 63 percent of the labor force sample and, as it is well known, estimates derived from self-selected samples may be biased due to correlations between the independent variables and the stochastic disturbance induced by the sample selection rule. I develop a two-step Heckman model in order to test whether this process is biasing the results, using the number of employed individuals in the household and the presence of children and elderly as exclusion restrictions. In Table 7, I report the estimates of the wage equation under different controls for self-selection into the labor force. Two features are noteworthy form these results. First, the inverse mills ratio is not significant in any specification. Second, a comparison with the results depicted in Table 6 does not suggest important differences between the point estimates. Thus, we can say with confidence that there is no evidence of self-selection biases in any of these two specifications.

Table 7: Effects of Informal Referrals on Hourly Wages (log), Sample Selection, Two-Stage Heckman, LSS 2003.

	Heckman 2 Stage	
	Spec 1 Wage Eq. $\hat{\gamma}_{X_1}$ [R-SE]	Spec 2 Wage Eq. $\hat{\gamma}_{X_2}$ [R-SE]
γ_0	-0.101 [0.029]***	-0.111 [0.023]***
Inverse Mill's Ratio	0.077 [0.086]	0.038 [0.087]
Constant	6.904 [0.091]***	6.923 [0.090]***
Control Variables	X_1	X_2
χ^2	13376	13810
N	18049	18049

NOTE: The table reports the coefficients of hourly wages in logarithms from a two-stage Heckman procedure. Here I only show the results the wage equation, for the results including the selection equation and other covariates refer to Appendix D table 13. The exclusion restrictions for this method are those variables that can affect the probability of being employed but not affect directly wages, I select those included in the employment equation: household variables (presence of children and elderly), number of employed individuals within the household, and whether the closest-relative used help from friends or relatives to obtain his actual job.

In sum, the results suggest there is no bias due to selection on unobservables nor sample-selection. Moreover, once we control for all relevant individual attributes, the result indicates that referred workers earn 13 percent less than non referred workers on average. From the theoretical literature we could argue that this result arises from the fact that informal referrals create poor matches: the availability of informal referrals and the opportunity to find a job more easily may induce a job seeker to accept a job offer in professions, sectors, or locations where his abilities are not fully exploited. Such a theory postulates a negative relationship between the use of informal referrals and wages because job seekers are willing to sacrifice higher wages to obtain a position rapidly. Therefore, using informal referrals may produce a mismatch between worker's comparative advantage and his occupational choices (Bentolila *et al.* 2004). However, this conclusion can be premature since the informal referrals dummy might be proxying for unobserved job attributes. For instance, informal sector vacancies are normally filled through informal referrals, and workers in this sector are more likely to earn less than workers in the formal sector on average. Given the endogeneity of individual employment attributes and wages, we cannot directly include these variables in the wage equation and, therefore, IR's coefficient might reflect the correlation between working in the informal sector and wages.

Thus, it is necessary to explore whether the IR effect differs between informal and formal sector workers. Here, I define informal sector workers as employees of small-size firms (i.e., less than 50 employees) without a signed contract nor social security benefits. Kugler (2003) is the only paper, to my knowledge, that tries to include a wage differential across sectors between referred and non referred workers. Unfortunately, a shortcoming of her paper is that it attempts to estimate wage differentials by using one or more dummy variables to indicate the sector where the individual is employed. However, if the labor market is segmented into formal and informal sectors, and one of them is more likely to hire workers through informal referrals, there might exist wage differentials across sectors that are not solved by including sector dummies or interaction effects. Moreover, ignoring the endogeneity of being in one sector or the other may bias the estimates that are based on sector-specific samples.

I propose to employ a model similar to the one commonly used to study union/non-union wage differentials (see Lee 1978, and Robinson & Tomes 1984), and public/private wage differentials (see Adamchik & Bedi 2000, Tansel 2005, and Van der Gaag & Vijverberg 1988). This model is known as the Roy model, or switching regression model. It consists of two wage equations and a "switching" equation that determines the sector in which the employee works (see Appendix C for a detailed description of the methodology). The wage equations include the same set of covariates used for the total sample; the switching equation (the probability of obtaining an informal-sector job) depends on the use of informal referrals IR, on individual attributes that are used by the employer to choose a worker (age, level of education, gender, race), the employment status of the closest relative, and, in Specification 2, the employment attributes of the closest relative.

In order to identify the parameters of the wage equation, we must impose the restriction that there is at least one regressor that directly affects preference for sector of employment but does not directly enter the wage equation (exclusion restriction). Here, I assume that the number of informal workers in the household, the presence of both children and elderly, and the informal referrals usage, directly affect the preferences of being employed in a given sector, but do not directly enter the wage equation. The motivation for using these exclusion restrictions is that individuals might use this information to decide in which sector to work and this information will only affect wages through the selection of the sector. For instance, the number of informal-sector workers in the household is a proxy variable for unobserved attributes that might be correlated at the household level. The presence of children and elderly is likely to influence the reservation wage, but unlikely to influence the gross offered wage and hence should only be included in the selection equation (Puhani, 2000).[18] In order to control for heterogeneous effects of the presence of children in the household by gender, I also include interaction terms of the presence of children with a dummy for males.

Table 8. Effect of Using Informal Referrals on Hourly Wages (log), Roy model, LSS 2003.

	FILM			Heckman 2 Stage		
	Wage Eq.	Wage Eq.	Sel Eq.	Wage Eq.	Wage Eq.	Sel Eq.
	Formal	Informal		Formal	Informal	
	b/se	b/se	b/se	b/se	b/se	b/se
Informal Referrals	-0.131	-0.076	0.668	-0.143	-0.077	0.689
	[0.013]***	[0.155]	[0.036]***	[0.021]***	[0.126]	[0.037]***
Constant	7.261	7.115	-0.933	7.221	6.863	-0.573
	[0.035]***	[0.439]***	[0.088]***	[0.060]***	[0.291]***	[0.095]***
Control Variables	X_1	X_1	X_3	X_1	X_1	X_3
ρ_0^a	-0.526					
	[.035]***					
ρ_1^b	-0.041					
	[.506]					
$Mill's Ratio$				0.091	-0.026	
				[0.045]**	[0.233]	

* p<.10, ** p<.05, *** p<.01
X_1 = [Male, White, Age, Education, Region, CR employment status]
X_3 = [Male, White, Age, Education, Region, CR employment status, Exclusion Restrictions]
NOTE: The table reports the coefficients of hourly wages in logarithms. The first three columns reports the results from a Full Information Maximum Likelihood, whereas, the last three columns show the results from the Heckman two stage methodology. For methodological description refer to Appendix C. The results for the selection equation show the coefficients from a probit model of the probability of being employed in the informal sector, they are coefficients and not marginal effects. a. ρ_0 is the correlation coefficient between the error term of the informal-sector wage equation, ϵ_{1i}, and the error term from the switching equation, u_i. b. ρ_1 is the correlation coefficient between the error term of the formal-sector wage equation, ϵ_{2i}, and the error term from the switching equation, u_i. See Appendix D Table 14, for the results with the entire set of covariates.

Table 8 shows the results obtained by two main estimation methods: FILM and Heckman two stages; refer to Appendix C for methodological details. For each method I present the three equations: the first two columns show the wage equations for each sector, and the third presents the results for the selection equation (i.e., the probability of working in the informal sector).

[18] Usual criticisms associated with these variables, such as household attributes having an impact on the tax rate, inducing a correlation between children and the after-tax wage, do not apply in this case because I am analyzing before-tax wages.

From the selection equation, it is evident that the use of informal referrals increases the likelihood of being employed in the informal sector. This result might arise from a number of reasons. First, informal sector firms are not constrained by labor regulations. Second, informal referrals lower hiring cost and are more likely to operate locally (Holzer, 1987a; Montgomery, 1991); thus, the informal sector firms are more likely to use informal referrals intensively than formal sector firms, not only because this might reduce cost but also because informal firms have better knowledge of the local network.

Notice that although workers using informal referrals might have better chances of obtaining a position in informal sector firms, there is no significant wage gap between referred and non referred workers. On the contrary, referred workers in formal-sector firms earn 13 percent less on average than non referred workers. This result might arise from the fact that jobs found through informal referrals in this specific sector are obtained more quickly but also pay lower wages, since at least some of them are filled by workers who sacrifice their productivity advantage in order to get a job more easily. This result is very important for the analysis of the effect of job search strategies on wages, since it confirms the mismatch theory between workers' productivity and occupational choices proposed by Bentolila *et al.* (2004). However, it is important to highlight that this is only true for formal-sector positions.

6. Conclusion

Theory predicts that job information networks play a relevant role on the flows of information between job seekers and potential employers. The reduction of informational asymmetries might facilitate employment, since individuals who use informal referrals to seek employment receive and accept more job offers than workers who use other methods. However, job seekers – at least partly – decide whether to use this method, and this determination may be correlated with both observed and unobserved characteristics which might also affect the probability of being employed. Moreover, both job seekers and potential employers might rely strongly on contacts to either find a job or fill vacancies, and this reliance arises from the institutional background instead of plausible amelioration of information flows between the two parties. Therefore, when drawing conclusions about causal effects of the use of informal referrals on employment one must take into account the possible confounding influence of other variables.

To explore for the presence of confounding influence of other variables I compare the average treatment effect on the employment status for those who use informal referrals as their main method to search for a job with those that use other "formal" methods.

In order to do this I use three main methodologies: ordinary least squares, propensity-score matching and instrumental variables. The first two assume that the confounding influence of other variables arises from those variables that we can observe in the data (individual and household specific attributes). When controlling for these attributes, either by OLS or PSM, the results suggest that the use of informal referrals increases the probability of being employed by 23 percentage points. There could exist, however, unobserved variables affecting this result such as communicative skills or institutional background. To test for their presence I carried out a two stage least squares estimation using the use of informal referrals from the closest relative in the household, conditional on the employment attributes of the closest relative, as an instrument. The results suggest that much of the positive effect stems from the confounding variables; I argue that the positive effect reflects the prevalence of informal-sector jobs to be filled through this method rather than a causal effect. Specifically, the presence of heterogeneous effects for those workers that are more likely to be hired in informal-sector jobs (males, young, and low educated) confirms this argument.

Several studies have evaluated whether the job positions acquired by informal channels lead to a differential in wages between referred and non referred workers. It is well known that existing theoretical and empirical literature has produced conflicting results. To explore how job information networks might affect the wage setting process in the Colombian labor market, I compare the results from OLS, PSM and IV. In this case, the results do not support the hypothesis of presence of unobserved confounding variables; on the contrary, they support the hypothesis that, conditional on observed individual attributes, referred and non referred workers become comparable.

The point estimates suggest that referred workers earn 13 percent less than non referred workers on average. Notice, however, that this result might be biased if either workers sort themselves into sectors, or if a worker's contacts have a positive impact on the probability of obtaining a job in, say, the informal sector, and this is correlated with the wage obtained in that sector. Given the endogeneity of employment attributes and wages, I use a Roy model to account for plausible wage differentials between informal and formal sectors. The "returns" to informal referrals are statistically different in both sectors: in the formal sector, referred workers earn less than non referred workers on average, while no difference is observed in the informal sector. Thus, we can conclude that informal referrals create low/poor-quality matches only in formal sector positions, since the availability of informal referrals and the opportunity to find a job more easily may induce a job seeker to accept positions in occupations where his abilities are not fully exploited.

References

Adamchik, V.A., & Bedi, A.S. 2000. Wage differentials between the public and the private sectors: Evidence from an economy in transition. *Labour economics*, **7**(2), 203–224.

Aguilera, M.B. 2002. The Impact of Social Capital on Labor Force Participation: Evidence from the 2000 Social Capital Benchmark Survey. *Social science quarterly*, **83**(3), 853–874.

Alon, S., & Stier, H. 2004. Job search, gender, and the quality of employment in israel. *Stratification in israel: Class, ethnicity, and gender.*

Amuedo-Dorantes, C., & Mundra, K. 2005 (Feb.). *Social networks and their impact on the employment and earnings of mexican immigrants*. Labor and Demography 0502001. EconWPA.

Angrist, J.D. 2001. Estimation of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice. *Journal of business & economic statistics*, **19**(1), 2–28.

Becker, S.O., & Ichino, A. 2002. Estimation of average treatment effects based on propensity scores. *The stata journal*, **2**(4), 358–377.

Bentolila, S., Michelacci, C., & Suarez, J. 2004. *Social contacts and occupational choice*. CEPR Discussion Paper No. 4308.

Bian, Y. 1997. Bringing Strong Ties Back in: Indirect Ties, Network Bridges, and Job Searches in China. *American sociological review*, **62**(3), 366–385.

Blau, D.M., & Robins, P.K. 1990. Job search outcomes for the employed and unemployed. *The journal of political economy*, **98**(3), 637–655.

Blundell, R., Dearden, L., & Sianesi, B. 2005. Evaluating the effect of education on earnings: models, methods and results from the National Child Development Survey. *Journal of the royal statistical society series a*, **168**(3), 473–512.

Boorman, S.A. 1975. A Combinatiorial Optimization Model for Transmission of Job Information Through Contact Networks. *The bell journal of economics*, **6**(1), 216–249.

Borghans, L., Weel, B., & Weinberg, B. 2006. *People people: Social capital and the labor market outcomes of underrepresented groups*. NBER Working Paper.

Bortnick, S.M., & Ports, M.H. 1992. Job Search Methods and Results: Tracking the Unemployed, 1991. *Monthly labor review*, **115**(12).

Bridges, W.P., & Villemez, W.J. 1986. Informal Hiring and Income in the Labor Market. *American sociological review*, **51**(4), 574–582.

Calvó-Armengol, A. 2004. Job contact networks. *Journal of economic theory*, **115**(1), 191–206.

Calvó-Armengol, A., & Jackson, M.O. 2004. The Effects of Social Networks on Employment and Inequality. *American economic review*, **94**(3), 426–454.

Calvó-Armengol, A., & Jackson, M.O. 2006. Networks in Labor Markets: Wage and Employment Dynamics and Inequality. *Journal of economic theory.*

Calvó-Armengol, A., & Zenou, Y. 2005. Job matching, social network and word-of-mouth communication. *Journal of urban economics*, **57**(3), 500–522.

Cameron, A.C., & Trivedi, P.K. 2005. *Microeconometrics: methods and applications*. Cambridge University Press.

Card, D. 1999. The causal effect of education on earnings. *Handbooks in economics*, 1801–1864.

Chapple, K. 2002. Out of Touch, Out of Bounds: How Social Networks Shape the Labor Market Radious of Women Welfare in San Francisco. *Urban geography*, **22**, 617–640.

Corcoran, M., Datcher, L., & Duncan, G. 1980b. Information and Influence Networks in Labor Markets. *Five thousand american families: Patterns of economic progress*, **8**(S 1), 37.

Datcher Loury, L. 2006. Some contacts are more equal than others: Informal networks, job tenure, and wages. *Journal of labour economics*, **24**(2), 299–318.

Elliot, J.R. 1999. Social isolation and labor market insulation: Network and neighborhood effects on less-educated urban workers. *Sociological quarterly*, **40**(2), 199–216.

Fernandez, R.M., & Castilla, E. 2001. How Much Is That Network Worth? Social Capital in Employee Referral Networks. *Social capital: Theory and research*, 85–104.

Fontaine, F. 2004 (June). *Do workers really benefit from their social networks?* Cahiers de la Maison des Sciences Economiques v04085. Université Panthéon-Sorbonne (Paris 1). available at http://ideas.repec.org/p/mse/wpsorb/v04085.html.

Goldberg, M.S. 1982. Discrimination, Nepotism, and Long-Run Wage Differentials. *The quarterly journal of economics*, **97**(2), 307–319.

Goos, M., & Salomons, A. 2007. Dangerous Liaisons: A Social Network Model for the Gender Wage Gap. *Center for economic studies.*

Granovetter, M. 1995. *Getting a Job: A Study of Contacts and Careers.* Chicago University Press.

Granovetter, M.S. 1973. The Strength of Weak Ties. *The american journal of sociology*, **78**(6), 1360–1380.

Green, GP, Tigges, LM, & Diaz, D. 1999. Racial and ethnic differences in job-search strategies in Atlanta, Boston, and Los Angeles. *Social science quarterly*, **80**(2), 263–278.

Heckman, J. 1997. Instrumental Variables: A Study of Implicit Behavioral Assumptions in One Widely Used Estimator. *Journal of human resorces*, **32**(3).

Heckman, J., Ichimura, H., & Todd, P.E. 1997. Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The review of economic studies*, **64**(4), 605–654.

Heckman, J., Ichimura, H., Smith, J., & Todd, P. 1998. Characterizing Selection Bias Using Experimental Data. *Nber working paper*, **6699**.

Heckman, J., LaLonde, R., & Smith, J. 1999. The Economics and Econometrics of Active Labor Market Policies. *Handbook of labor economics*, **3**, 1865–2097.

Holzer, H.J. 1987a. *Hiring Procedures in the Firm: Their Economic Determinants and Outcomes.*

Holzer, H.J. 1987b. Job Search by Employed and Unemployed Youth. *Industrial and labor relations review*, **40**(4), 601–611.

Imbens, G. W., & Angrist, J. D. 1994. Identification and estimation of local average treatment effects. *Econometrica*, **62**(2), 467–475.

Ioannides, Y.M., & Loury, L.D. 2004. Job Information Networks, Neighborhood Effects, and Inequality. *Journal of economic literature*, **42**(4), 1056–1093.

Korpi, T. 2001. Good friends in bad times? social networks and job search among the unemployed in sweden. *Acta sociologica*, **44**(2), 157–170.

Kugler, A.D. 2003. Employee Referrals and Efficiency Wages. *Labour economics*, **10**(5), 531–556.

Labini, M.S. 2004. *Social Networks and Wages: It is all about connections!* Tech. rept. LEM Working Paper Series.

Lee, L.F. 1978. Unionism and wage rates: a simultaneous equations model with qualitative and limited dependent variables. *International economic review*, 415–433.

Lin, N. 2001. Building a Network Theory of Social Capital. *Social capital: Theory and research.*

Marmaros, D., & Sacerdote, B. 2002. Peer and social networks in job search. *European economic review*, **46**(4-5), 870–879.

Marsden, P.V., & Gorman, E.H. 2001. *Social Networks, Job Changes, and Recruitment.* Springer.

Montgomery, J.D. 1991. Social Networks and Labor-Market Outcomes: Toward an Economic Analysis. *The american economic review*, **81**(5), 1408–1418.

Mortensen, Dale T., & Vishwanath, Tara. 1994. Personal contacts and earnings : It is who you know! *Labour economics*, **1**(2), 187–201.

Mouw, T. 2003. Social Capital and Finding a Job: Do Contacts Matter? *American sociological review*, **68**(6), 868–898.

Osberg, L. 1993. Fishing in Different Pools: Job-Search Strategies and Job-Finding Success in Canada in the Early 1980s. *Journal of labor economics*, **11**(2), 348–386.

Pellizzari, M. 2004. *Do Friends and Relatives Really Help in Getting a Good Job?* London School of Economics and Political Science-Centre for Economic Performance.

Puhani, P. 2000. The Heckman correction for sample selection and its critique. *Journal of economic surveys*, **14**(1), 53–68.

Robinson, C., & Tomes, N. 1984. Union wage differentials in the public and private sectors: a simultaneous equations specification. *Journal of labor economics*, 106–127.

Rosenbaum, J.E., DeLuca, S., Miller, S.R., & Roy, K. 1999. Pathways into Work: Short-and Long-Term Effects of Personal and Institutional Ties. *Sociology of education*, **72**(3), 179–196.

Rosenbaum, P.R., & Rubin, D.B. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika*, **70**(1), 41–55.

Simon, C.J., & Warner, J.T. 1992. Matchmaker, Matchmaker: The Effect of Old Boy Networks on Job Match Quality, Earnings, and Tenure. *Journal of labor economics*, **10**(3), 306–330.

Tansel, A. 2005. Public-private employment choice, wage differentials, and gender in Turkey. *Economic development and cultural change*, **53**(2), 453–477.

Todd, P. 2006. Evaluating Social Programs with Endogenous Program Placement and Selection of the Treated. *Handbook of development economics*, **4**.

Van der Gaag, J., & Vijverberg, W. 1988. A switching regression model for wage determinants in the public and private sectors of a developing country. *The review of economics and statistics*, 244–252.

Wooldridge, J.M. 2003. Further results on instrumental variables estimation of average treatment effects in the correlated random coefficient model. *Economics letters*, **79**(2), 185–191.

Yakubovich, Valery. 2005. Weak ties, information, and influence: How workers find jobs in a local russian labor market. *American sociological review*, **70**, 408–421(14).

Appendix A. Data Description

The Living Standards Survey (LSS) provides measures of the socioeconomic status of the Colombian population. In particular, it provides information on school enrolment rates, housing conditions, access to amenities and facilities, income and expenditures, unemployment rates, health indicators, and child care among others. The LSS is representative at the regional level (five regions and four main departments: Atlantic, Eastern, Central, Pacific, Orinoquia, Antioquia, Valle del Cauca, San Andres and Bogota).

Sampling methodology: The sample for the LSS survey is taken from the universe of the census population. The sampling methodology consist of first generating strata according to geographical location and socioeconomic level; then randomly drawing municipalities from these strata; next, randomly drawing neighborhoods from these municipalities; and finally, randomly drawing blocks and then households from these neighborhoods. To facilitate the collection of information, households are grouped into segments of ten households on average.

Main Variables used in the Analysis:

Variable	Description
Informal Referrals (IR)	Persons who during the reference period stated one of the following categories: i. are seeking for a job through friends or relatives (unemployed), ii. obtained their current job trough friends or relatives (employed).
Employment	Persons between 15 and 65 years old who during the reference period were in one of the following categories: *i.* persons who worked for at least one hour for remuneration during the last week, *ii.* persons who did not work during the reference week but had a job.
Monthly Wages	Wages are composed of remuneration in cash in the primary and secondary job.
Hourly Wages	The monthly wage divided by the monthly working hours.
Age Groups	Set of dummy variables: less than 15, 15-24, 25-45, 45-64, 65 and more (the reference age group is 15-24).
Level of Education	Refers to the highest level of attained education. It is measured with a set of dummy variables, one for each educational level, which in each case equal 1 if the person was in one of the following categories: *i.* Primary or Less: 0 to 5 years of education. *ii.* Secondary: 6 to 11 years of education. *iii.*Non-university higher education: bachelors degree of three years *iv.* University: bachelors degree of five years. This also includes postgraduate education.
White	Describes the ethnic group of each member of the household. It is a dummy variable equal to 1 if the individual declares himself as white.
Marital Status	Describes the marital status of each member of the household in the moment of the survey. In the analysis a dummy variable equal to 1 if the individual is single was included.
Region	Refers to the geographic area where the household is located. Set of dummy variables: Atlantico, Central, Pacific, Eastern, Orinoquia- Amazonia, Antioquia, Valle del Cauca, San Andres, and Bogota (reference group).
Closest Blood Relative (CR)	It is a variable that indicates the closest person to each individual within the household. For example, for offspring the closest member might be the father or mother; in their absence, the closest blood relative can be selected among the stepfather or stepmother, the grandfather or grandmother, a sibling, uncle, cousin, and so forth. It has to respect this order. However, for the head of the household who is married or cohabiting this variable takes the information of his or her partner. For uni-personal households this variable takes the value of zero
CR Employed	It is a binary variable equal to 1 if the closest blood relative is employed.
CR employed by IR	It is a binary variable equal to 1 if the closest blood-relative is employed and found his current job through the help of friends or relatives.
CR employment attributes	They are a set of variables that describe the employment attributes of the closest relative.

Appendix B. Interaction Effects

Wooldridge (2003) shows that standard instrumental variable estimators, applied to an equation containing interactions, consistently estimate the average treatment effect. Under some assumptions, we can estimate the following model in two stages:

$$y = x\beta + \gamma IR + x_1 IR\delta_1 + x_2 IR\delta_2 \ldots + x_g IR\delta_g + u$$

First estimate a linear reduced form for IR by regressing IR on $(1, x, z)$ and obtaining the fitted values, say $\hat{IR}$. Valid IVs for observation i are then $(1, x_i, \hat{IR}_i, x_{1i}\hat{IR}_i, x_{2i}\hat{IR}_i, ..., x_{gi}\hat{IR}_i)$. Mechanically, one would specify the variables from the previous equation and list as IVs those in brackets. The interpretation of the coefficients, however, is not straightforward. Say we want to test if the use of informal referrals is associated with an increase in the probability of being employed when condition x_j is met. The marginal effect of using informal referrals, evaluated at the reference value of all other variables, is given by $\partial y_i / \partial \hat{IR} = \gamma + \delta_1 x_j$ for $j = 1, 2, ..., g$. From here we see that γ only captures the effect of IR on employment when all conditioning variables are set to zero, i.e. $x_1 = x_2 = \cdots = x_w = 0$. Similarly, the coefficient β_j only captures the effect of each variable on employment for non IR users, i.e. $IR = 0$. Thus, it is incorrect to say that a positive and significant coefficient on IR (or x_j) indicates that a variation in IR (or x_j) is expected to lead to an increase in the probability of employment. Moreover, the effect of using informal referrals on employment depends on the values of the conditioning variables. If the conditioning variable is, for example, gender we should present two values: the marginal effect of IR when x_1 is zero and when x_1 is one, i.e., γ, and $\gamma + \delta_1$, along with the corresponding standard errors. The standard errors of interest are: $\sigma_{y,\hat{IR}} = \sqrt{var(\hat{\gamma}) + x_j^2 var(\hat{\delta_j}) + 2x_j cov(\hat{\gamma}, \hat{\delta_j})}$. We can also estimate the average treatment effect on the treated as: $ATT = \sum_{i=1}^{N}(\hat{\gamma} + x_{i1}\hat{\delta}_1 + x_{i2}\hat{\delta}_2 + \ldots + x_{ig}\hat{\delta}_g)$, and clearly one should also estimate the correct standard errors, i.e., $\sigma_{ATT} = \sqrt{var(\hat{\gamma}) + \sum_j^g x_j^2 var(\hat{\delta_j}) + \sum_j^g 2x_j cov(\hat{\gamma}, \hat{\delta_j})}$. These are the results depicted in Table 5.

Appendix C. The Roy Model

Let us assume that a criterion function I_i determines in which sector the agent is employed. $I_i = 1$ if worker i is employed in the informal sector, and $I_i = 0$ if he is employed in the formal sector:

$$I_i = 1 \quad \text{if} \quad \gamma Z_i + u_i > 0$$
$$I_i = 0 \quad \text{if} \quad \gamma Z_i + u_i \leq 0$$

Now, we can assume that hourly wages in each sector are determined as follows:

$$\begin{aligned} ln(w_{1i}) &= \beta_1 X_{1i} + \epsilon_{1i} \qquad if\, I_i = 1 \\ ln(w_{2i}) &= \beta_2 X_{2i} + \epsilon_{2i} \qquad if\, I_i = 0 \end{aligned}$$

Where w_{ji} is the hourly wage and X_{ji} is a vector of wage determining variables (e.g. gender, race, age, and level of education). Assuming normality of u, ϵ_1 and ϵ_2, maximum likelihood estimates of the parameter vectors of interest, β_1 and β_2, can be obtained. For identification this methodology relies on two assumptions. First, there is at least one variable that affects the decision to work in a given sector but does not directly hourly wages. In this set up the variables that seem to satisfy this condition are the following: number of informal workers in the household, presence of both children and elderly individuals, and a dummy variable if the closest relative in the household obtained his job by informal referrals. Second, it assumes that u, ϵ_1 and ϵ_2 have a trivariate normal distribution with mean zero (in vector form) and covariance matrix:

$$\sum = \begin{bmatrix} \sigma_u^2 & . & . \\ \sigma_{21} & \sigma_1^2 & . \\ \sigma_{31} & . & \sigma_2^2 \end{bmatrix}$$

where σ_u^2 is the variance of the error term in the selection equation, and σ_1^2 and σ_2^2 are the variances of the error terms in the wage equations. σ_{21} is the covariance of u_1 and ϵ_{1i} and σ_{31} is the covariance of u_i and ϵ_{2i}. The covariance between ϵ_{1i} and ϵ_{2i} is not defined as w_{1i} and w_{2i} are never observed simultaneously. Given the assumption of the error terms the parameters can be estimated through Full Information Maximum Likelihood (FILM). It is more common, however, to estimate the model using Heckman's two step method applied to the truncated means (Cameron & Trivedi, 2005):

$$\begin{aligned} E[ln(w_i)|X_i, I_i = 1] &= \beta_1 X_1 + \sigma_{21}\lambda(\gamma Z) \\ E[ln(w_i)|X_i, I_i = 1] &= \beta_2 X_2 - \sigma_{31}\lambda(-\gamma Z) \end{aligned}$$

Where $\lambda(\gamma Z)$ is the inverse mills ratio, i.e., $\lambda(\gamma Z) = \phi(\gamma Z)/\Phi(\gamma Z)$. A first-stage probit estimation of whether or not worker i works in the informal sector yields an estimate of γ and hence $\lambda(\gamma Z)$. Two separate OLS regressions then lead to direct estimates of (β_1, σ_{21}) and (β_2, σ_{31}). In the text I report the estimates from both the full information maximum likelihood and the Heckman two step method applied to truncated means.

Appendix D. Tables

Table 9. Effect of using informal referrals on the probability of being employed, OLS and 2SLS, LSS 2003

	SPECIFICATION 1			SPECIFICATION 2		
	OLS	2SLS		OLS	2SLS	
	Empl. Eq	Empl. Eq	IR Eq.	Empl. Eq	Empl. Eq	IR Eq.
Constant	0.412	0.393	0.655	0.418	0.535	0.65
	[0.020]***	[0.057]***	[0.023]***	[0.020]***	[0.058]***	[0.023]***
Informal Referrals	0.232	0.26		0.23	0.051	
	[0.006]***	[0.082]***		[0.006]***	[0.083]	
Individual Variables						
Male dummy	0.029	0.028	0.033	0.026	0.032	0.033
	[0.007]***	[0.008]***	[0.009]***	[0.007]***	[0.008]***	[0.009]***
White dummy	0.049	0.047	0.044	0.049	0.057	0.044
	[0.014]***	[0.014]***	[0.017]***	[0.014]***	[0.014]***	[0.017]***
Single dummy	-0.041	-0.041	-0.016	-0.037	-0.04	-0.017
	[0.007]***	[0.008]***	[0.009]*	[0.007]***	[0.008]***	[0.010]*
Age 25-34	0.143	0.142	0.023	0.144	0.148	0.024
	[0.009]***	[0.009]***	[0.010]**	[0.009]***	[0.009]***	[0.010]**
Age 35-44	0.151	0.149	0.064	0.152	0.163	0.066
	[0.010]***	[0.011]***	[0.012]***	[0.010]***	[0.011]***	[0.012]***
Age 45-54	0.139	0.138	0.044	0.14	0.149	0.048
	[0.011]***	[0.012]***	[0.014]***	[0.011]***	[0.012]***	[0.014]***
Age 55-64	0.087	0.086	0.041	0.088	0.095	0.044
	[0.017]***	[0.017]***	[0.020]**	[0.017]***	[0.017]***	[0.020]**
Secondary	0.053	0.056	-0.092	0.056	0.041	-0.088
	[0.009]***	[0.012]***	[0.010]***	[0.009]***	[0.012]***	[0.010]***
High educ non university	0.122	0.127	-0.189	0.128	0.096	-0.179
	[0.011]***	[0.019]***	[0.014]***	[0.011]***	[0.019]***	[0.014]***
Unfinished university	0.106	0.112	-0.186	0.117	0.086	-0.171
	[0.012]***	[0.020]***	[0.015]***	[0.012]***	[0.019]***	[0.015]***
University	0.163	0.17	-0.242	0.175	0.134	-0.226
	[0.011]***	[0.023]***	[0.014]***	[0.011]***	[0.022]***	[0.014]***
Post University	0.239	0.249	-0.359	0.254	0.192	-0.336
	[0.011]***	[0.032]***	[0.016]***	[0.012]***	[0.031]***	[0.017]***
Household Variables						
Child 0 -3 years old	-0.044	-0.044	-0.012	-0.044	-0.046	-0.008
	[0.014]***	[0.014]***	[0.016]	[0.014]***	[0.014]***	[0.016]
Child 4 -6 years old	-0.003	-0.003	0.016	-0.002	0.001	0.018
	[0.013]	[0.013]	[0.016]	[0.013]	[0.013]	[0.016]
Child 6 - 12 years old	-0.023	-0.023	-0.003	-0.022	-0.022	-0.002
	[0.010]**	[0.010]**	[0.012]	[0.010]**	[0.010]**	[0.012]
Child 0 -3 years old x Male	0.101	0.1	0.042	0.097	0.103	0.038
	[0.016]***	[0.016]***	[0.021]**	[0.016]***	[0.017]***	[0.021]*
Child 4 -6 years old x Male	0.048	0.049	-0.03	0.044	0.039	-0.034
	[0.016]***	[0.016]***	[0.021]	[0.016]***	[0.016]**	[0.021]
Child 6 -12 years old x Male	0.04	0.04	0.017	0.038	0.04	0.015
	[0.013]***	[0.013]***	[0.016]	[0.013]***	[0.013]***	[0.016]
Presence old individual	-0.021	-0.021	0.003	-0.018	-0.018	0.001
	[0.010]**	[0.010]**	[0.012]	[0.010]*	[0.010]*	[0.012]
Closest Relative (CR) Variables						
CR Type of employment						
Self-employed				-0.033	-0.041	0.024
				[0.009]***	[0.010]***	[0.013]*
CR Sector						
Public Sector				-0.04	-0.063	-0.101
				[0.015]***	[0.019]***	[0.020]***
CR Firm Size						
Medium size				0.001	-0.003	-0.027
				[0.013]	[0.013]	[0.017]
Large Size				0.011	0.005	-0.035
				[0.011]	[0.012]	[0.015]**
CR other						
White Collar employee				0.011	0.013	0.018
				[0.009]	[0.009]	[0.010]*
CR earnings (logs)				-0.018	-0.019	-0.006
				[0.004]***	[0.004]***	[0.005]
CR employed	0.001	0.001	-0.023	0.027	0.035	-0.029
	[0.006]	[0.006]	[0.009]***	[0.011]**	[0.012]***	[0.015]*
CR inactive	0.057	0.056	0.036	0.042	0.047	0.028
	[0.012]***	[0.013]***	[0.018]**	[0.013]***	[0.013]***	[0.019]
Unipersonal Household	0.029	0.029	-0.008	0.017	0.015	-0.013
	[0.015]*	[0.015]**	[0.019]	[0.015]	[0.015]	[0.019]
Excluded Instrument						
CR employed by IR			0.092			0.114
			[0.010]***			[0.012]***
Other Controls						
Region dummies	YES	YES	YES	YES	YES	YES
R-squared	0.148	0.147	0.054	0.15	0.106	0.058
N	18049	18049	18049	18049	18049	18049
Endogeneity test		0.123			4.938	
Endogeneity p-value		0.726			0.026	
Weak instruments Test		91.34			89.655	

* p<.10, ** p<.05, *** p<.01

Table 10. Effect of using informal referrals on the probability of being employed, Interaction Effects, LSS 2003

	SPECIFICATION 1 IV + Interactions		SPECIFICATION 2 IV + Interactions	
	Emp. Eq	IR x Var	Emp. Eq	IR x Var
Constant	0.181		0.422	
	[0.165]		[0.153]***	
Informal Referrals	0.535		0.207	
	[0.236]**		[0.220]	
Individual Variables				
Male dummy	-0.089	0.195	-0.076	0.18
	[0.034]***	[0.055]***	[0.031]**	[0.051]***
White dummy	-0.021	0.106	-0.007	0.1
	[0.063]	[0.103]	[0.061]	[0.100]
Single dummy	-0.041		-0.039	
	[0.008]***		[0.008]***	
Age 25-34	0.243	-0.168	0.249	-0.172
	[0.058]***	[0.094]*	[0.055]***	[0.089]*
Age 35-44	0.284	-0.217	0.314	-0.245
	[0.061]***	[0.097]**	[0.057]***	[0.091]***
Age 45-54	0.31	-0.276	0.347	-0.321
	[0.062]***	[0.100]***	[0.059]***	[0.094]***
Age 55-64	0.382	-0.465	0.39	-0.467
	[0.079]***	[0.127]***	[0.078]***	[0.123]***
Attained Education				
Secondary	0.186	-0.153	0.033	0.029
	[0.133]	[0.174]	[0.123]	[0.161]
High educ non university	0.423	-0.419	0.299	-0.31
	[0.158]***	[0.230]*	[0.141]**	[0.202]
Unfinished university	0.588	-0.738	0.299	-0.322
	[0.165]***	[0.245]***	[0.147]**	[0.214]
University	0.457	-0.422	0.351	-0.359
	[0.148]***	[0.213]**	[0.133]***	[0.188]*
Post University	0.683	-0.842	0.504	-0.676
	[0.150]***	[0.239]***	[0.132]***	[0.200]***
Household Variables				
Child 0 -3 years old	-0.043		-0.045	
	[0.014]***		[0.014]***	
Child 4 -6 years old	0.001		0.003	
	[0.014]		[0.014]	
Child 6 - 12 years old	-0.021		-0.02	
	[0.011]**		[0.011]*	
Child 0 -3 years old x Male	0.092		0.095	
	[0.017]***		[0.017]***	
Child 4 -6 years old x Male	0.045		0.035	
	[0.016]***		[0.017]**	
Child 6 -12 years old x Male	0.032		0.033	
	[0.013]**		[0.013]**	
Presence old individual	-0.022		-0.019	
	[0.010]**		[0.010]*	
Closest Relative (CR) Variables				
CR Type of employment				
Self-employed			-0.038	
			[0.010]***	
CR Sector				
Public Sector			-0.078	
			[0.019]***	
CR Firm Size				
Medium size			0.001	
			[0.013]	
Large Size			0.009	
			[0.012]	
CR other				
White Collar employee			0.012	
			[0.009]	
CR earnings (logs)			-0.02	
			[0.004]***	
CR employed	0.001		0.034	
	[0.007]		[0.012]***	
CR inactive	0.059		0.05	
	[0.013]***		[0.013]***	
Unipersonal Household	0.032		0.016	
	[0.015]**		[0.015]	
Other Controls				
Region Dummies	YES		YES	
R-squared	0.076		0.08	
N	18049		18049	

* p<.10, ** p<.05, *** p<.01

Table 11. Effect of using informal referrals on hourly wages, OLS and 2SLS, LSS 2003

	SPECIFICATION 1			SPECIFICATION 2		
	OLS	2SLS		OLS	2SLS	
	Wage. Eq	Wage. Eq	IR Eq.	Wage. Eq	Wage. Eq	IR Eq.
Constant	7.304	7.307	0.714	7.27	7.265	0.708
	[0.034]***	[0.116]***	[0.026]***	[0.034]***	[0.076]***	[0.026]***
Informal Referrals	-0.134	-0.138		-0.131	-0.124	
	[0.012]***	[0.156]		[0.012]***	[0.098]	
Individual Variables						
Male dummy	0.129	0.129	0.024	0.139	0.138	0.026
	[0.011]***	[0.011]***	[0.009]***	[0.011]***	[0.011]***	[0.009]***
White dummy	-0.016	-0.015	0.036	-0.015	-0.015	0.033
	[0.027]	[0.028]	[0.022]	[0.027]	[0.027]	[0.021]
Age 25-34	0.281	0.28	-0.042	0.277	0.277	-0.038
	[0.015]***	[0.017]***	[0.012]***	[0.015]***	[0.016]***	[0.012]***
Age 35-44	0.408	0.407	-0.021	0.405	0.405	-0.017
	[0.016]***	[0.017]***	[0.013]	[0.016]***	[0.016]***	[0.013]
Age 45-54	0.559	0.559	-0.031	0.555	0.555	-0.024
	[0.020]***	[0.020]***	[0.015]**	[0.020]***	[0.020]***	[0.015]
Age 55-64	0.56	0.56	-0.034	0.556	0.556	-0.029
	[0.033]***	[0.034]***	[0.023]	[0.033]***	[0.033]***	[0.023]
Attained Education						
Secondary	-0.338	-0.337	0.081	-0.32	-0.32	0.078
	[0.016]***	[0.020]***	[0.012]***	[0.016]***	[0.017]***	[0.012]***
High educ non university	0.397	0.397	-0.123	0.37	0.371	-0.117
	[0.018]***	[0.026]***	[0.015]***	[0.018]***	[0.021]***	[0.015]***
Unfinished university	0.513	0.513	-0.125	0.475	0.476	-0.123
	[0.019]***	[0.028]***	[0.016]***	[0.019]***	[0.023]***	[0.016]***
University	0.988	0.987	-0.19	0.928	0.929	-0.182
	[0.019]***	[0.037]***	[0.015]***	[0.020]***	[0.027]***	[0.015]***
Post University	1.362	1.36	-0.318	1.284	1.287	-0.304
	[0.022]***	[0.055]***	[0.019]***	[0.023]***	[0.039]***	[0.019]***
Closest Relative (CR) Variables						
CR Type of employment						
Self-employed				0.057	0.057	0.148
				[0.018]***	[0.018]***	[0.024]***
CR Sector						
Public Sector				-0.076	-0.075	-0.185
				[0.033]**	[0.033]**	[0.031]***
CR Firm Size						
Medium size				0.017	0.017	0.019
				[0.025]	[0.025]	[0.020]
Large Size				0.093	0.093	-0.008
				[0.023]***	[0.023]***	[0.019]
CR other						
White Collar employee				0.084	0.084	0.01
				[0.017]***	[0.017]***	[0.013]
CR earnings (logs)				0.071	0.071	0.004
				[0.008]***	[0.008]***	[0.006]
CR employed	-0.024	-0.024	-0.016	-0.143	-0.144	0.006
	[0.012]**	[0.012]**	[0.011]	[0.020]***	[0.020]***	[0.016]
CR inactive	0.047	0.048	0.02	0.102	0.102	0.02
	[0.028]*	[0.028]*	[0.021]	[0.029]***	[0.029]***	[0.022]
Unipersonal Household	-0.064	-0.064	-0.036	-0.023	-0.022	-0.032
	[0.026]**	[0.026]**	[0.022]	[0.026]	[0.026]	[0.022]
Excluded Instrument						
CR employed by IR			0.099			0.215
			[0.013]***			[0.023]***
Other Controls						
Region dummies	YES	YES	YES	YES	YES	YES
R-squared	0.511	0.511	0.071	0.519	0.519	0.083
N	11468	11468	11468	11468	11468	11468
Endogeneity test		0.001			0.005	
Endogeneity p-value		0.977			0.943	
Weak instruments Test		53.732			85.437	

* p<.10, ** p<.05, *** p<.01

Table 12. Effect of using informal referrals on hourly wages, interaction effects, LSS 2003

	SPECIFICATION 1 Interactions		SPECIFICATION 2 Interactions	
	Wage. Eq	IR x Var	Wage. Eq	IR x Var
Constant	6.897		6.892	
	[0.062]***		[0.062]***	
Informal Referrals	-0.141		-0.14	
	[0.012]***		[0.012]***	
Individual Variables				
Male dummy	0.137	-0.011	0.144	-0.008
	[0.018]***	[0.022]	[0.018]***	[0.023]
White dummy	0.002	-0.027	0.001	-0.025
	[0.043]	[0.054]	[0.043]	[0.054]
Age 25-34	0.271	0.011	0.264	0.015
	[0.028]***	[0.033]	[0.028]***	[0.033]
Age 35-44	0.414	-0.012	0.409	-0.01
	[0.030]***	[0.036]	[0.030]***	[0.035]
Age 45-54	0.553	0.007	0.554	-0.003
	[0.034]***	[0.041]	[0.034]***	[0.041]
Age 55-64	0.652	-0.132	0.645	-0.129
	[0.049]***	[0.065]**	[0.050]***	[0.065]**
Attained Education				
Secondary	0.393	-0.073	0.375	-0.073
	[0.036]***	[0.039]*	[0.035]***	[0.039]*
High educ non university	0.773	-0.051	0.731	-0.059
	[0.041]***	[0.049]	[0.041]***	[0.048]
Unfinished university	0.921	-0.101	0.874	-0.116
	[0.043]***	[0.051]**	[0.042]***	[0.050]**
University	1.338	-0.002	1.266	-0.015
	[0.039]***	[0.049]	[0.040]***	[0.049]
Post University	1.685	0.069	1.598	0.053
	[0.041]***	[0.053]	[0.042]***	[0.055]
Closest Relative (CR) Variables				
CR Type of employment				
Self-employed			0.019	
			[0.031]	
CR Sector				
Public Sector			-0.039	
			[0.048]	
CR Firm Size				
Medium size			-0.058	
			[0.043]	
Large Size			0.053	
			[0.037]	
CR other				
White Collar employee			0.101	
			[0.029]***	
CR earnings (logs)			0.065	
			[0.012]***	
CR employed	-0.018		-0.179	
	[0.020]		[0.034]***	
CR inactive	0.077		0.132	
	[0.043]*		[0.044]***	
Unipersonal Household	-0.1		-0.057	
	[0.043]**		[0.043]	
Other Controls				
Region Dummies	YES		YES	
R-squared	0.512		0.522	
N	11468		11468	

* p<.10, ** p<.05, *** p<.01

Table 13. Effect of using informal referrals on hourly wages, Two Stage Heckman, LSS 2003

	Specification 1		Specification 2	
	Wage Eq.	Select Eq.	Wage Eq.	Select Eq.
	b/[sd]	b/[sd]	b/[sd]	b/[sd]
Constant	6.904	-0.652	6.923	-0.634
	[0.091]***	[0.071]***	[0.090]***	[0.071]***
Informal Referrals	-0.1	0.777	-0.11	0.775
	[0.029]***	[0.023]***	[0.030]***	[0.023]***
Individual Variables				
Male dummy	0.142	0.106	0.146	0.099
	[0.014]***	[0.028]***	[0.013]***	[0.028]***
White dummy	0.005	0.162	0.004	0.165
	[0.027]	[0.052]***	[0.027]	[0.052]***
Age 25-34	0.312	0.552	0.298	0.55
	[0.026]***	[0.030]***	[0.026]***	[0.030]***
Age 35-44	0.443	0.645	0.43	0.642
	[0.029]***	[0.033]***	[0.029]***	[0.034]***
Age 45-54	0.6	0.598	0.589	0.592
	[0.029]***	[0.039]***	[0.029]***	[0.039]***
Age 55-64	0.619	0.386	0.611	0.378
	[0.033]***	[0.058]***	[0.033]***	[0.058]***
Attained Education				
Secondary	0.33	0.212	0.312	0.22
	[0.018]***	[0.034]***	[0.018]***	[0.034]***
High educ non university	0.743	0.447	0.695	0.461
	[0.026]***	[0.046]***	[0.026]***	[0.046]***
Unfinished university	0.874	0.36	0.815	0.384
	[0.025]***	[0.046]***	[0.026]***	[0.047]***
University	1.379	0.557	1.299	0.58
	[0.027]***	[0.046]***	[0.028]***	[0.048]***
Post University	1.827	0.955	1.726	0.981
	[0.036]***	[0.060]***	[0.038]***	[0.062]***
Closest Relative (CR) Variables				
CR Type of employment				
Self-employed			0.063	-0.122
			[0.017]***	[0.041]***
CR Sector				
Public Sector			-0.007	-0.138
			[0.030]	[0.062]**
CR Firm Size				
Medium size			0.016	0.035
			[0.026]	[0.054]
Large Size			0.075	0.01
			[0.022]***	[0.046]
CR other				
White Collar employee			0.078	0.059
			[0.017]***	[0.032]*
CR earnings (logs)			0.073	-0.039
			[0.007]***	[0.014]***
CR employed	-0.03	-0.019	-0.209	0.055
	[0.012]**	[0.030]	[0.020]***	[0.049]
CR inactive	0.058	0.214	0.113	0.187
	[0.026]**	[0.062]***	[0.026]***	[0.063]***
Unipersonal Household	-0.047	0.12	-0.007	0.089
	[0.028]*	[0.057]**	[0.028]	[0.058]
Excluded Instrument				
CR informal		0.029		0.032
		[0.012]**		[0.013]**
CR employed by IR		0.013		-0.054
		[0.031]		[0.038]
Children 0 3		-0.066		-0.067
		[0.048]		[0.048]
Children 4 6		0.062		0.06
		[0.049]		[0.049]
Children 7 12		-0.05		-0.045
		[0.038]		[0.038]
Children 0 3 x male		0.368		0.355
		[0.068]***		[0.068]***
Children 4 6 x male		0.213		0.204
		[0.071]***		[0.071]***
Children 7 12 x male		0.168		0.161
		[0.053]***		[0.053]***
Old individuald		-0.137		-0.131
		[0.036]***		[0.037]***
Ancilliary Parameters				
Mills Ratio	0.077		0.038	
	[0.086]		[0.087]	
Other Controls				
Region dummies	YES	YES	YES	YES
Model chi-square	13376.41		13810.817	
N	18049		18049	

* p<.10, ** p<.05, *** p<.01

Table 14. Effect of using informal referrals on hourly wages, Switching Regression, LSS 2003

	FILM			Heckman 2 Stage		
	Wage Informal b/se	Wage Formal b/se	Sel. Eq b/se	Wage Informal b/se	Wage Formal b/se	Sel. Eq b/se
Constant	7.115	7.261	-0.933	6.863	7.22	-0.573
	[0.439]***	[0.035]***	[0.088]***	[0.291]***	[0.060]***	[0.095]***
Informal Referrals	-0.076	-0.131	0.668	-0.077	-0.143	0.689
	[0.155]	[0.013]***	[0.036]***	[0.126]	[0.021]***	[0.037]***
Individual Variables						
Male dummy	0.218	0.106	0.064	0.218	0.111	0.088
	[0.025]***	[0.011]***	[0.037]*	[0.026]***	[0.011]***	[0.038]**
White dummy	-0.058	0.001	0.063	-0.058	0.003	0.049
	[0.057]	[0.027]	[0.068]	[0.056]	[0.027]	[0.069]
Age 25-34	0.273	0.264	-0.418	0.274	0.231	-0.411
	[0.088]***	[0.017]***	[0.039]***	[0.071]***	[0.022]***	[0.040]***
Age 35-44	0.306	0.407	-0.517	0.306	0.369	-0.492
	[0.112]***	[0.018]***	[0.043]***	[0.086]***	[0.024]***	[0.043]***
Age 45-54	0.363	0.569	-0.788	0.364	0.516	-0.756
	[0.176]**	[0.021]***	[0.052]***	[0.135]***	[0.031]***	[0.052]***
Age 55-64	0.282	0.614	-0.554	0.282	0.576	-0.525
	[0.133]**	[0.033]***	[0.080]***	[0.107]***	[0.036]***	[0.080]***
Attained Education						
Secondary	0.256	0.348	-0.41	0.256	0.312	-0.405
	[0.089]***	[0.020]***	[0.039]***	[0.071]***	[0.024]***	[0.039]***
High educ non university	0.211	0.417	-0.675	0.467	0.686	-1.055
	[0.164]	[0.018]***	[0.052]***	[0.189]**	[0.041]***	[0.060]***
Unfinished university	0.323	0.537	-0.635	0.58	0.808	-1.001
	[0.153]**	[0.020]***	[0.054]***	[0.178]***	[0.042]***	[0.062]***
University	0.587	1.016	-0.897	0.844	1.279	-1.257
	[0.230]**	[0.018]***	[0.061]***	[0.237]***	[0.043]***	[0.066]***
Post University	1.206	1.373	-1.285	1.464	1.629	-1.689
	[0.354]***	[0.022]***	[0.112]***	[0.349]***	[0.047]***	[0.115]***
Closest Relative (CR) Variables						
CR employed	-0.052	-0.012	0.074	-0.052	-0.005	0.076
	[0.033]	[0.012]	[0.038]**	[0.031]*	[0.012]	[0.039]**
CR inactive	-0.112	0.084	0.098	-0.113	0.092	0.152
	[0.066]*	[0.027]***	[0.077]	[0.067]*	[0.026]***	[0.077]**
Unipersonal Household	-0.028	-0.067	0.202	-0.029	-0.049	0.212
	[0.074]	[0.029]**	[0.071]***	[0.068]	[0.029]*	[0.072]***
Exclusion Restrictions						
Informal Workers HH			0.005			0.019
			[0.017]			[0.017]
CR employed by IR			0.001			0.003
			[0.038]			[0.039]
Children 0 3			0.187			0.178
			[0.066]***			[0.066]***
Children 4 6			0.147			0.149
			[0.064]**			[0.064]**
Children 6 12			0.03			0.03
			[0.050]			[0.053]
Children 0 3 x male			-0.139			-0.138
			[0.081]*			[0.081]*
Children 4 6 x male			-0.182			-0.187
			[0.080]**			[0.081]**
Childrem 6 12 x male			-0.021			-0.01
			[0.063]			[0.066]
Old individuals			-0.013			0.002
			[0.052]			[0.051]
Ancilliary Parameters						
Mills Ratio					0.09	-0.026
					[0.045]**	[0.233]
Lns	-0.526	-0.63				
	[0.020]***	[0.011]***				
rho1	-0.041	-0.585				
	[0.508]	[0.050]***				
Other Controls						
Region Dummies	YES	YES	YES	YES	YES	YES
Model chi-square	247.359			264.12	3739.787	
N	11468			11468	11468	

* p<.10, ** p<.05, *** p<.01

Institut de Recherches Économiques et Sociales
Université catholique de Louvain

Place Montesquieu, 3
1348 Louvain-la-Neuve, Belgique

ISSN 1379-244X D/2009/3082/023