

Why do migrant workers rely more often on referrals ?

Sevak Alaverdyan*

Abstract

Empirical analysis of the SOEP data from 2002 to 2008 show that 41.21% of migrant workers and 31.79% of native workers found their current job through referrals. Estimation results of the panel probit model with random effects show that a part of the 9.42% can be explained by the control variables including characteristics of the individuals and firms. But still there is 7.2% statistically significant difference of the predicted probabilities of using referrals between migrant and native workers. Intuitively, one can expect that migrant workers have smaller social networks in the new destination country and by that smaller probability of finding a job through referrals, but our empirical observation shows that migrant workers are more likely to find a job through referrals even after introducing controls. In order to explain this puzzle, this paper presents a search and matching model with heterogeneous worker groups and several search channels. Calibration results of the model show that even when migrant workers have smaller size of the social network, they gain more from recommendations, because their employment chances are initially much lower than employment chances of native workers.

Keywords: social contacts, referrals, productivity signal, matching function

JEL Classification: J15, J24, J64, J71

*Bielefeld Graduate School of Economics and Business, Bielefeld University, 33615 Bielefeld, Germany

1 Introduction

A large literature has examined the methods used by workers and firms in the job matching process both theoretically and empirically. Earlier studies were conducted by sociologists to emphasize the importance of the social contacts when observing search of the firms and workers. Informal and formal search methods of firms and workers are classified in the literature in the following way: formal search methods include search through newspaper advertisements, state and private employment agencies, school and college placement services and etc, informal methods include search through relatives, friends, acquaintances, referrals from other employees and etc.

Different models were developed and empirical studies were conducted to show the effect of search methods of firms and workers on the labour market outcomes. Let us first discuss the models and the main findings of the literature, particularly, the model predictions and empirical results on the effects of search methods on the wage and probability to be hired. Most of the studies conducted agree that using informal search methods increases the probability to be hired, but the model predictions and empirical results on wage effect are mixed: some authors find positive wage effect (see Montgomery (1991), Dustmann et al. (2011), Galenianos (2013)), while others find the effect to be negative (Pistaferri (1999), Bentolila et al. (2010)).

According to Rees (1966) and Marsden (1988) social contacts tend to occur among workers with similar characteristics. Rees (1966) further claims that a worker will refer only well-qualified applicants, since his reputation is at stake. Following these arguments, Montgomery (1991) constructs an adverse-selection model with two time periods, where employer observes the type of a worker and conditional on the observed type of the current worker makes a wage offer for the next period. Thus employers relying on referrals from high ability workers try to mitigate the adverse-selection problem, assuming that the current high ability worker will refer to an own type high ability worker. Dustmann et al. (2011) build their model on the Jovanovic (1979, 1984) job matching model, and they extend the model by distinguishing between informal and formal search methods. The key difference between the two search methods according to them is that the worker's match-specific productivity is more uncertain when using formal methods, rather than informal methods. Overall, based on their assumptions both models predict positive effect of using informal search methods on the match quality between firms and workers.

While Bentolila et al. (2010) propose a model, according to which matching through informal channels produces mismatch between the worker's productivity and occupation. In their model they assume that every worker has productive advantage in exactly one occupation and the social contacts of the worker are all employed in same occupation, which is random. So there is a positive probability that the use of informal methods by the worker may cause mismatch and thus lower quality matches between workers and firms.

Unlike the model predictions on the wages, the predictions on the probability to be hired are similar for most of the above mentioned models. Dustmann et al. (2011) do not directly explore the probability to be hired when using different search channels. While Galenianos (2013) uses similar learning model and predicts positive effect of using referrals on the probability to be hired. The model by Montgomery (1991) also predicts positive effect of using referrals on the probability to be hired. The positive effect can be explained by the expectations of the firm,

that most likely high ability worker will refer to a worker of his/her own type. According to the favoritism explanation, referred workers are more likely to be hired due to the influence by the referee. Although Bentolila et al. (2010) find mismatch between worker's most productive and actual occupation, they find positive effect on probability to obtain the job when using referrals.

Besides theoretical studies of the effect of using referrals on the match quality there are empirical observations in the literature. Brown, Setren and Topa (2014) exploit panel dataset on a single U.S corporation and empirically check the model predictions in the literature. Using this dataset enables them to control for various individual and job specific characteristics, but as the data is on a single firm, the results can hardly be representative for the entire economy. They obtain 2.4% positive effect of using referrals on probability to be hired; moreover, conditional on being interviewed the positive effect is 13.9%.

Cappellari and Tatsiramos (2013) draw attention to the findings of Calvo-Armengol and Jackson (2004) that the key factor for understanding the effectiveness of the social contacts is employment status of the contacts. Following this idea instead of using proxies of network size, they use British Household Panel Survey, which provides the employment status of the closest three friends, which they define as the network quality measure. They estimate the effect of the network quality on the job finding probability using three estimation methods to eliminate potential bias and obtained positive relationship between the number of employed friends and job finding probability. Cappellari and Tatsiramos (2013) propose an alternative explanation to this result, arguing that the higher network quality of network can make pressure to non-employed member to search more actively leading to higher job finding rate.

Drever and Hoffmeister (2008) use the SOEP data and find that nearly half of all the migrants find their job through networks. In addition, they find that when the migrants find the job through social networks the perceived improvement of working conditions is the same as if they find the job through formal channels. Moreover, the improvement of working conditions does not depend on the ethnic makeup of the migrant's network. While Lancee (2016) finds that for migrants using referrals leads to higher earnings than using formal search channels only in case of the bridging social capital, but the result holds only for high educated migrants with good German language proficiency. Eisnecker and Schacht (2016) observe the length of time it took to find the first job for migrants and find that the migrants, which use informal search channels find the job faster than the migrants which search through formal channels. These studies report the effect of using referrals on the labour market outcomes. But this study rather aims to find the reason of the different frequency of using referrals between native and migrant workers.

Empirical analysis of the SOEP data from 2002 to 2008 show that 41.21% of migrant workers found their current job through referrals. While 31.79% of native workers found their current job through referrals. Estimation results of the panel probit model with random effects show that a part of the difference can be explained by the control variables including characteristics of the individuals and firms. But still there is a statistically significant difference of the predicted probabilities of using referrals between migrant and native workers. Intuitively, one can expect that migrant workers have smaller social networks in the new destination country and by that smaller probability finding a job through referrals, but our empirical observation shows that migrant workers are more likely to find a job through referrals even after introducing controls.

The main goal of this paper is to analyze this paradox and to explain why migrant workers find jobs through social contacts more often than natives even if they are expected to do it less frequently than natives.

This paper presents a search and matching model with heterogeneous worker groups and several search channels. In the model the firm does not observe the real productivity of the worker, the firm observes the productivity signal of the worker, the group and matching channel of the worker. There are four groups of workers: natives (immigrants) who found the vacancy through referrals or formal search channels. We assume that compared to the workers without referral, recommended workers produce different signal to the firm. Furthermore, we assume that natives' and migrants' productivity signals are different either. Phelps (1973) suggests two cases where two individuals from different groups and with the same signal can be treated differently. First, when the average productivity of the groups differs, and the signals are equally informative. Second, when the average productivity of the groups are the same, while the signals are differently informative. Dustmann et al. (2011) and Galenianos (2013) assume different uncertainty of signals between the workers matched through referrals and formal channels, which is like the second case. We rather follow the first case when introducing inequality between the groups.

The worker is hired if the expected productivity given the productivity signal is higher than the lower bound of the productivity for the group of the worker. The lower bounds of the groups are determined based on the ex-ante bargained wage. Calibration of the model gives the distributions of the productivity signals for the four groups. Thus, we can obtain the probability of being hired after the match for the four groups. Even though in the model native workers have more social contacts than migrants, migrant workers are still more likely to find a job through referrals. Results from the calibration of the model show that the average productivity difference between the native and migrant workers is positive. Moreover, average productivity difference between the workers with and without referrals is also positive. The probability that the native workers are hired after they are matched with the firm through referrals is approximately 2.15 times more than the probability for the native workers matched through formal channels. The probability that the migrant workers are hired after they are matched with the firm through referrals is 4.06 times more than the probability for the migrant workers matched through formal channels. So the gain from finding job through referrals for the migrants is higher than the gain for the natives.

The study proceeds as follows: section 2 describes the data used, the descriptive statistics and discusses the empirical approach and estimation results; section 3 explains the search and matching model with heterogeneous worker groups and several search channels; section 4 presents the results from the calibration of the model; section 5 concludes.

2 Empirical Analysis

2.1 Data

This subsection describes the data used and defines the dependent and independent variables. This study uses data from the German Socio-Economic Panel(SOEP) for the empirical analysis.

The German Socio-Economic Panel every year covers nearly 11,000 households, and about 30,000 individuals. SOEP is a longitudinal study of households and individuals. Among a wide range of questions regarding personal characteristics and employment data respondents are asked how they found their current job. Information about the way the respondents found their current job makes SOEP data suitable for this study. Our sample covers data on 6769 employed individuals from SOEP 2002-2008.

Individuals are considered to have found job through referrals if they responded that they found out about their current job through friends or relatives. Individuals are considered to have found job through formal channels in case they responded that they found out about their current job through other channels, for example, through the federal employment office, an advertisement in the internet or in the newspaper, through a job-center (ARGE) and etc. The corresponding dummy variable REF_{it} takes value 1, if the i^{th} person found out about the job through referrals at time t , and it takes value 0, if the i^{th} person found out about the job through formal channels at time t . Another categorical variable $TOJCH_{it}$ indicates which kind of job change finding the current job was. $TOJCH_{it}$ has 5 categories: first job, job after break, job with new employer, company taken over, changed job at the same firm. $FSIZE_{it}$ is another categorical variable, which shows the size of the firm the i^{th} individual is employed at time t . It has 6 categories: less than 5 employees, 5 to 19, 20 to 99, 100 to 199, 200 to 1999, and 2000 and more employees.

One of the most important variables is the variable MIG_{it} indicating the nationality of individuals. We define an individual to be migrant if the person has foreign citizenship or got German citizenship at a later date than birth. And the German or native are individuals who got German citizenship at their birth. The variable MIG_{it} equals to 1 if the i^{th} individual is migrant at time t , and it is equal to 0 if the i^{th} individual is German at time t . As a measure of individual's education we use the amount of education or training in years computed by the SOEP.¹ The corresponding variable EDU_{it} shows the i^{th} individual's computed education or training in years at time t . The values of EDU_{it} range from 7 to 18. AGE_{it} shows the i^{th} individual's age at time t . In our analysis we consider 18 to 65 year old individuals.²

To control for the occupational status we used the Standard International Socio-Economic Index of Occupational Status developed by Ganzeboom et al. (1992). Based on information about individual's income, education, and occupation, ISEI index reflects individual's socio-economic status. $ISEI_{it}$ equals to the i^{th} individual's ISEI value at time t and ranges from 16 to 90. The last variable to be introduced is the $LTWF_{it}$, which shows the length of time with the firm i^{th} individual worked at the firm at the t point of time.

2.2 Descriptive Statistics

In this subsection we discuss the descriptive statistics of the variables used in the estimations. REF_{it} is the main variable of our interest. According to the descriptive statistics presented in the Table 2, 32.9% of observed respondents replied, that they found their current job through referrals. On the one hand, according to the Table 1, there is significant 9.42% difference between the native and migrant individuals which found the current job through referrals. On the other

¹for detailed description see Helberger (1988) and Schwarze (1991).

² AGE_{it} is equal to the difference of the year of survey minus year of the individual's birth.

hand, 44.40% of foreign citizens found the current job through referrals, compared to 32.04% for citizens of Germany. So in this case the difference is 12.36%, which shows that for the migrants with German citizenship the percentage is closer to the one for natives, compared to the percentage for migrants with foreign citizenship. This might possibly be explained by some sort of assimilation of migrants with German citizenship.

Citizenship\Migration status	Found job through referrals
Foreign Citizens	44.40%
Citizens of Germany	32.04%
Migrants	41.21%
Natives	31.79%

Table 1: Percentage of individuals found the current job through referrals by citizenship\migration status.

Table 2 presents descriptive statistics of the variables used in the estimations. Besides the statistics about the overall sample observed, Table 2 includes descriptive statistics separately for the natives and the migrants, to better understand the differences between the two groups.

Variable	Natives	Migrants	Overall
EDU_{it}	12.69	11.35	12.53
AGE_{it}	34.364	34.358	34.36
AGE_{it}^2	1300.56	1282.35	1282.35
REF_{it}	0.318	0.412	0.329
MIG_{it}	-	-	0.1217
$FSIZE_{it}$	%	%	%
less than 5	11.43	13.00	11.62
5 to 19	22.52	23.70	22.67
20 to 99	21.01	23.36	21.29
100 to 199	8.83	8.41	8.78
200 to 1999	17.26	16.57	17.18
2000 and more	18.95	14.95	18.46
$TOJCH_{it}$	%	%	%
first job	12.08	11.05	11.95
job after break	32.92	40.10	33.80
job with new employer	42.86	42.14	42.77
company taken over	2.58	2.55	2.57
changed job same firm	9.56	4.16	8.90
$LTWF_{it}$	1.75	1.21	1.69
$ISEI_{it}$	44.83	38.38	44.04

Table 2: Descriptive statistics

On average migrants have 11.35 years of education, which is 1.34 years less compared to the average education of natives. Moreover, migrants are relatively younger and are employed in smaller firms compared to natives. When we look at the first three categories of $FSIZE_{it}$, we can see that proportion of migrants employed in each of the categories is higher than the proportion of natives employed at the same category. On the contrary, higher proportion of natives is employed in each of the last three categories of $FSIZE_{it}$ in comparison to the proportion of migrants employed at the same category. When it comes to the type of job change, except from

the category job after "break" in all the other categories proportion of migrants is lower. In the case of "job after break" and "changed job same firm," the differences are relatively higher. In the former case the proportion is higher for the migrants, in the latter case, for the natives. Furthermore, length of time with the firm is 1.75 years for the natives, while for the migrants this indicator is 1.21, so the migrants in our sample have less experience of working at the same firm compared to natives. Likewise, average $ISEI_{it}$ is higher for the natives, so compared to migrants, natives have higher average occupational status. To sum up, in our sample migrants are relatively younger, with lower average years of education, with lower average ISEI index, are less experienced at the current firm and are employed at smaller-size firms.

2.3 Empirical Approach

Since the dependent variable $REF_{it} = 1$ is binary, we use binary choice regression models in empirical estimations. Moreover, since the data used is longitudinal, we compare the results obtained by using pooled probit regressions and panel probit regressions with random effects. Among different regression models these two regression models were chosen and we discuss the reasons in this subsection. Let us first discuss fixed effects logit and probit models for panel data. When number of the time periods T is fixed, estimation of the fixed effects model encounters an incidental parameters problem. As a result, estimators of the constant terms are not consistent. And since the maximum-likelihood estimator (MLE) of coefficients is a function of the estimators of the constant terms, MLE of the coefficients is inconsistent either. (see Neyman and Scott(1948) and Lancaster (2000).) However, following Rasch (1960) and Andersen (1970), Chamberlain (1980) proposes a consistent conditional maximum likelihood estimator (CMLE) given that the conditional likelihood function satisfies regularity conditions. These regularity conditions impose mild restrictions on the incidental parameters discussed in Andersen (1970),(1971). Chamberlain (1980) demonstrates that conditional on sufficient statistics for the incidental parameters, likelihood function is free of the incidental parameters. In the logit model, sum of the individual dependent variable's value over time is a minimal sufficient statistic for the individual constant term. Thus, the CMLE is computationally convenient estimator for the fixed effects logit model, but it is not the case for the fixed effects probit model. In the fixed effects probit model the incidental parameters can not be removed from the conditional likelihood function, because there are no sufficient statistics available for the probit model. Hence, in case of fixed effects the logit model is more preferable than the probit model.

However, conditional ML estimation of fixed effects logit model is not efficient for our estimations for the following reasons. First, out of 9670 observations only 2376 are used in the CMLE estimations, because if the individual dependent variable does not change over time, the conditional probability of the observation contributes nothing to the conditional likelihood function. Second, the marginal effects can not be estimated with the coefficients estimated by CMLE, because fixed effects are not estimated.

Unlike the case of the fixed effects models, for the analysis of the random effects, the probit model is more preferable than the logit model. Since in the logit model errors are assumed to have a logistic distribution, logit model uses multivariate logistic distribution. The disadvantage of the multivariate logistic distribution is that the correlations are all constrained to be 0.5. Thus

the probit model, which is based on the multivariate normal distribution, is more flexible. (see Johnson and Kotz (1972), Maddala (1987).)

Compared to the binary choice fixed effects model, disadvantage of the binary choice random effects models is that these models do not allow for a correlation between the individual effects and the explanatory variables. Random effects probit model produces a consistent estimator of coefficients under the certain very strong assumptions about the heterogeneity. (see Greene (2007), section 23.5.) Random effects model can be extended to binary choice setting by the method specified by Butler and Moffitt (1982). Then, log likelihood can be approximated using a Gauss-Hermite quadrature technique. Estimation of the random effects probit model was conducted using the statistical program Stata, which follows adaptive Gauss-Hermite quadrature method of Naylor and Smith (1982) to approximate the panel-level likelihood.³

We compare the estimation results of the random effects probit model to the results obtained using pooled probit model. Our observations show that there is not statistically significant difference between using pooled logit or pooled probit model, so we choose to use a pooled probit model since the results are then better comparable to the results obtained from the estimation of random effects probit model.⁴

2.4 Estimation Results

First, we compare the panel probit estimator to the pooled probit estimator. A likelihood-ratio test is conducted to check if the panel estimator is different from the pooled estimator. The test suggests that the panel-level variance component is significantly more than zero, which implies that the panel probit model with random effects is statistically more preferable than the pooled probit model. Table 3 shows the estimated coefficients and marginal effects at mean values of the variables both for the pooled probit model and panel probit model with random effects. The estimated coefficients and marginal effects of the two models are not very different from each other. This can be explained by the fact, that approximately only 1.4 observations are available on average per individual. However, estimation results of the panel probit RE will be used for the further analysis.

The positive coefficient of the variable MIG_{it} shows that migrants are more likely to be hired through referrals after the controls. After the controls the difference of 9.42% decreases to 6.2% when we compare a migrant with a native worker both with the same average characteristics. Negative coefficients of variables EDU_{it} and AGE_{it} show that individuals with higher age and more education are less likely to find the current job through referrals. Moreover, individuals occupied in the firms with more employees are less likely to find job through referrals. The probability of finding job through referrals of a migrant with average characteristics ranges from 42.5% for the firm size category "less than 5" to 27.8% for the firm size category "2000 and more." For a native with average characteristics the corresponding probabilities are 35.7% and 22.2% respectively.

$TOJCH_{it}$ is yet another significant variable. The probability of finding job through referrals

³See details at the Stata 13 Base Reference Manual.

⁴Furthermore, there are other alternative estimation techniques proposed in the literature. But estimators from both the pooled probit model and random effects probit model give consistent estimators. Moreover, estimating these two models is computationally convenient and fits better to the data particularly we have.

of an individual with average characteristics is 29.1% when the current job of the individual is the first job. The probability is not significantly different when the job is a job after break. The corresponding probability equals to 27.5%. But in case of job with new employer the probability is significantly different and equals to 33.5%. Unlike the three cases mentioned above, the probability is relatively low in the cases of company taken over and changed job same firm. In the former case the probability is 10.8%, in the latter case it is 12.8%.

Dustmann et al. (2011), Brown et al. (2012), Kugler (2003) and Galenianos (2013) find in their empirical and theoretical studies that the separation rates are lower for the individuals which found their job through referrals. Thus one would expect the coefficient of $LTWF_{it}$ to be positive, but it is negative. Even though it is counter intuitive, it does not contradict to the findings mentioned above. In the SOEP survey individuals mention the length of time with the firm if they obtained the job during the previous two years. So the negative coefficient might be a result of a selection. And finally, an individual with average characteristics is less likely to find job through referrals if the individual's $ISEI_{it}$ is higher.

Variable	Pooled probit Coefficient	Pooled probit Marginal effects	Panel probit RE Coefficient	Panel probit RE Marginal effects
Const	0.721***		0.737***	
EDU_{it}	-0.019***	-0.0067***	-0.021***	-0.0070***
AGE_{it}	-0.024***	-0.0011*	-0.026***	-0.0012*
AGE_{it}^2	0.00031**		0.00033**	
MIG_{it}	0.172***	0.061***	0.178***	0.062***
$FSIZE_{it}$				
5 to 19	-0.124***	-0.046***	-0.136***	-0.050***
20 to 99	-0.217***	-0.079***	-0.237***	-0.085***
100 to 199	-0.232***	-0.084***	-0.248***	-0.088***
200 to 1999	-0.324***	-0.115***	-0.352***	-0.122***
2000 and more	-0.368***	-0.129***	-0.400***	-0.137***
$TOJCH_{it}$				
job after break	-0.053	-0.018	-0.049	-0.017
job with new employer	0.107**	0.039**	0.124**	0.044**
company taken over	-0.634***	-0.180***	-0.686***	-0.183***
changed job same firm	-0.546***	-0.161***	-0.585***	-0.163***
$ISEI_{it}$	-0.0051***	-0.0017***	-0.0056***	-0.0019***

Significance level: * if $p < 0.10$, ** if $p < 0.05$, *** if $p < 0.01$

Table 3: Estimated coefficients and marginal effects.

To sum up, a part of the different frequency of finding job through referrals between natives and immigrants can be explained by individual characteristics and other controls. However, there are different predicted probabilities which may be used for the numerical analysis. Table 4 shows different predicted probabilities using pooled probit model and panel probit RE.

"Migrants/natives at means" are two otherwise-average individuals' predicted probabilities of using referrals. "Average migrant/native" is the predicted probability of using referrals for an individual with average migrant/native features. "Average of migrants/natives" is the average of the predicted probabilities of using referrals for migrants/natives. In case of the "all migrants/natives" all observations are treated as if they are all migrant/native regardless of

	Predicted probabilities	
	Pooled probit	Panel probit RE
Migrants at means	0.351	0.335
Natives at means	0.289	0.273
Average migrant	0.391	0.378
Average native	0.284	0.268
Average of migrants	0.412	0.403
Average of natives	0.318	0.306
All migrants	0.382	0.371
All natives	0.322	0.310
Average migrant*	0.359	0.344
Average native*	0.288	0.272

Table 4: Predicted probabilities

their migration status. Then the probabilities are predicted using the features of all observations (treating them as migrant/native). The resulting averages are "all migrants/natives." "Average migrant*/native*" are the predicted probabilities of a migrant/native with average education and age of migrants/natives keeping other variables at the overall means.⁵ Since in the model migrants and natives have different average productivity, we will use the predicted probabilities of average migrant*/native* in the numerical example. We propose a model and bring a numerical example, where we try to find the possible reason for the 7.2% difference between "average migrant*" and "average native*."

3 Model

We develop a search and matching model with heterogeneous worker groups and several search channels. The model is built on the "Equilibrium Unemployment Theory" of Pissarides (2000). A continuum of risk neutral workers and firms live forever and discount future at a common discount rate r . Firms are homogeneous, and there is free entry of new vacancy with the flow cost c . In this setting the firm does not observe real productivity of the applicant, instead the firm observes a noisy signal of the productivity, as well as the nationality of the applicant. The firm also knows whether the applicant found the vacancy through referrals or through formal channels.

3.1 Productivity

There are two groups of workers, $j = n$ - natives and $j = i$ - migrants. The true productivity of workers p is normally distributed with mean μ and variance σ^2 . ϵ is a zero-mean error that is normally distributed with variance σ_ϵ^2 . This productivity is not observable by the firm.

$$p \sim N(\mu, \sigma^2), \epsilon \sim N(0, \sigma_\epsilon^2) \quad (1)$$

Unemployed workers are identical within their ethnic group j . Once the worker is matched with the firm, there is a match-specific productivity draw p . There are two search channels in

⁵All the probabilities are predicted assuming that the random effects are equal to 0.

the model: formal and informal (referrals). In the model the firm does not observe the real productivity of the worker, the firm observes the productivity signal of the worker, the group and matching channel of the worker. In the empirical part we find that in our sample on average migrants have less education. Moreover, Kaas and Manger (2012) conduct a field experiment on the ethnic discrimination in Germany's labour market and conclude that there might be statistical discrimination in hiring in Germany's labour market. So, we assume that natives' and migrants' productivity signals are different, and the difference equals to d . Thus, the the productivity draw after being matched for the native workers p_n and for the migrant workers p_i are respectively:

$$p_n \sim N(\mu, \sigma^2), p_i \sim N(\mu - d, \sigma^2) \quad (2)$$

Kugler (2003) finds that firms using informal methods in hiring, lower their monitoring cost, because referees exert peer pressure on the referred workers. This means that referrals raise the productivity of the recommended workers. This enables the firms to use referrals and pay lower efficiency wages. In equilibrium, the matching process generates segmentation in the labour market, where well-connected workers are matched through referrals to high paying jobs, while less-connected workers are matched through formal channels to lower paid jobs. So, the recommended workers are more productive compared to the workers matched through formal channels. Thus, we assume that compared to the workers without referral workers with referral produce different signal to the firm. On average the difference between the productivity signals is s . Based on the findings in the literature related to the job search through social networks and hiring discrimination we expect both s and d to be positive. And indeed we find them to be positive in the calibration results. So the productivity signal of the native worker without referral p'_{nw} is:

$$p'_{nw} = p + \epsilon, p'_{nw} \sim N(\mu, \sigma^2 + \sigma_\epsilon^2) \quad (3)$$

Productivity signal of the native worker with referral p'_{nc} is:

$$p'_{nc} = p + s + \epsilon, p'_{nc} \sim N(\mu + s, \sigma^2 + \sigma_\epsilon^2) \quad (4)$$

Productivity signal of the migrant worker without referral p'_{iw} is:

$$p'_{iw} = p - d + \epsilon, p'_{iw} \sim N(\mu - d, \sigma^2 + \sigma_\epsilon^2) \quad (5)$$

Productivity signal of the migrant worker with referral p'_{ic} is:

$$p'_{ic} = p - d + s + \epsilon, p'_{ic} \sim N(\mu - d + s, \sigma^2 + \sigma_\epsilon^2) \quad (6)$$

The firm forms unbiased expectations on the real productivity of applicants given the group of the worker and the noisy productivity signal. Since the true productivity of the workers and the productivity signals are jointly normally distributed, following DeGroot (2005) expected productivity given the productivity signal for the four groups can be written as follows. The

expected productivity given the productivity signal of the native worker without referral:

$$E(p|p'_{nw}) = \mu + \frac{\sigma^2}{\sigma^2 + \sigma_\epsilon^2}(p'_{nw} - \mu) \quad (7)$$

The expected productivity given the productivity signal of the native worker with referral:

$$E(p|p'_{nc}) = \mu + s + \frac{\sigma^2}{\sigma^2 + \sigma_\epsilon^2}(p'_{nc} - \mu - s) \quad (8)$$

The expected productivity given the productivity signal of the migrant worker without referral:

$$E(p|p'_{iw}) = \mu - d + \frac{\sigma^2}{\sigma^2 + \sigma_\epsilon^2}(p'_{iw} - \mu + d) \quad (9)$$

The expected productivity given the productivity signal of the migrant worker with referral:

$$E(p|p'_{nc}) = \mu - d + s + \frac{\sigma^2}{\sigma^2 + \sigma_\epsilon^2}(p'_{nc} - \mu + d - s) \quad (10)$$

3.2 Matching function

Since the workers are matched to the firm through formal channels and through referrals, the matching functions for the two cases are different. We assume that the matching functions for native and migrant workers $m_{fj}(u, v)$ have Cobb-Douglas form in case of the matching through formal channels.

$$m_{fn}(u, v) = \lambda_0 u_n^\eta v^{1-\eta}, m_{fi}(u, v) = \lambda_0 u_i^\eta v^{1-\eta} \quad (11)$$

Where u_j is the unemployment rate of the group j , v is the vacancy rate, and η is the elasticity with respect to the unemployment in the matching function. λ_{fn} and λ_{fi} are the job finding rates in the formal market for natives and immigrants respectively:

$$\lambda_{fn} = \frac{\lambda_0 v^{(1-\eta)}}{u_n^{(1-\eta)}}, \lambda_{fi} = \frac{\lambda_0 v^{(1-\eta)}}{u_i^{(1-\eta)}} \quad (12)$$

Where λ_0 is the total factor productivity of the matching function, u_n and u_i are the number of unemployed natives and migrants respectively. Job filling rates in the formal market for natives q_n and immigrants q_i :

$$q_n = \frac{\lambda_0 u_n^\eta}{v^\eta}, q_i = \frac{\lambda_0 u_i^\eta}{v^\eta} \quad (13)$$

Following Stupnytska and Zaharieva (2015), matching functions in case of matching through referrals for natives, m_n can be written as:

$$m_n = va((f - u_n)[1 - (1 - \frac{u_n}{f})^{\gamma_n h_n} (1 - \frac{u_i}{1-f})^{(1-\gamma_n)h_n}] \frac{\gamma_n \frac{u_n}{f}}{\gamma_n \frac{u_n}{f} + (1-\gamma_n) \frac{u_i}{1-f}} + (1-f-u_i)[1 - (1 - \frac{u_i}{1-f})^{\gamma_i h_i} (1 - \frac{u_n}{f})^{(1-\gamma_i)h_i}] \frac{(1-\gamma_i) \frac{u_n}{f}}{\gamma_i \frac{u_i}{1-f} + (1-\gamma_i) \frac{u_n}{f}}) \quad (14)$$

Where a is the exogenous rate at which a vacancy arrives to a worker per unit time. (see Cahuc and Fontaine (2009).) f is the fraction of natives, $1 - f$ is the fraction of migrants. γ_n is the level of homophily between the social contacts of a native worker. The corresponding rate of homophily of a migrant worker is γ_i . h_n is the number of social contacts of natives, and the number of social contacts of migrants is h_i . In this setup the match between the firm and a native worker happens in the following way. The firm opens v vacancies, and va vacancies arrive to the workers of the firm. Information about these vacancies can be transmitted to an unemployed native both by an employed native and by an employed migrant worker. $va(f - u_n)$ vacancies arrive to an employed native worker. $(1 - \frac{u_n}{f})^{\gamma_n h_n}$ is the probability that all the native contacts of the native worker are employed. $(1 - \frac{u_i}{1-f})^{(1-\gamma_n)h_n}$ is the probability that all the migrant contacts of the native worker are employed. So $[1 - (1 - \frac{u_i}{1-f})^{\gamma_i h_i} (1 - \frac{u_n}{f})^{(1-\gamma_i)h_i}]$ is the probability that the employed native worker has at least one unemployed contact. The employed native worker transmits the information to a random unemployed contact, which is unemployed native with probability $\frac{\gamma_n \frac{u_n}{f}}{\gamma_n \frac{u_n}{f} + (1 - \gamma_n) \frac{u_i}{1-f}}$. With probability $(1 - f - u_i)$ the va vacancies arrive to an employed migrant worker. The probability that the employed migrant worker has at least one unemployed contact is $[1 - (1 - \frac{u_i}{1-f})^{\gamma_i h_i} (1 - \frac{u_n}{f})^{(1-\gamma_i)h_i}]$. The employed migrant worker transmits the information to a random unemployed contact, which is unemployed native with probability $\frac{(1 - \gamma_n) \frac{u_i}{1-f}}{\gamma_n \frac{u_n}{f} + (1 - \gamma_n) \frac{u_i}{1-f}}$.

Following the same intuition and the vacancy information transmission mechanism described above the matching function in case of matching through referrals for migrants, m_i can be written as:

$$m_i = va((1 - f - u_i)[1 - (1 - \frac{u_i}{1-f})^{\gamma_i h_i} (1 - \frac{u_n}{f})^{(1-\gamma_i)h_i}] \frac{\gamma_i \frac{u_i}{1-f}}{\gamma_i \frac{u_i}{1-f} + (1 - \gamma_i) \frac{u_n}{f}} + (f - u_n)[1 - (1 - \frac{u_n}{f})^{\gamma_n h_n} (1 - \frac{u_i}{1-f})^{(1-\gamma_n)h_n}] \frac{(1 - \gamma_n) \frac{u_i}{1-f}}{\gamma_n \frac{u_n}{f} + (1 - \gamma_n) \frac{u_i}{1-f}}) \quad (15)$$

Job finding rate in case of finding job through referrals is λ_{nc} for natives and λ_{ic} for migrants:

$$\lambda_{nc} = \frac{m_n}{u_n}, \lambda_{ic} = \frac{m_i}{u_i} \quad (16)$$

3.3 Bellman equations

In the model the firm does not observe the real productivity of the workers, but it observes the group of the worker and the distribution of the productivity signal of that group. As the wages are set by an ex-ante bargaining between the union and the firm, the wage is set the same for all workers and does not depend on the productivity of an individual worker. So if we assume that the wage w is more than the unemployment benefit b ,⁶ then the Bellman equation for the present discounted value of an employed native worker W_n can be written as:

$$rW_n = w - \delta(W - U_n) \quad \text{or} \quad W_n - U_n = \frac{w - rU_n}{(r + \delta)} \quad (17)$$

⁶ $w > b$ condition is necessary to ensure that $\frac{w - rU_n}{(r + \delta)} > 0$

Where δ is the exogenous job destruction rate, and U_n is the present discounted value of unemployed native workers.

Bellman equation for the present discounted value of an employed migrant worker W_I can be written as:

$$rW_i = w - \delta(w - U_i) \text{ or } W_i - U_i = \frac{w - rU_i}{(r + \delta)} \quad (18)$$

Where U_i is the present discounted value of unemployed migrant workers.

Bellman equation for the present discounted value of unemployed native workers:

$$rU_n = b + \lambda_{fn} \int_{p'_{0nw}}^{\infty} [W_n - U_n] d\Phi_{nw}(p'_{nw}) + \frac{m_n}{u_n} \int_{p'_{0nc}}^{\infty} [W_n - U_n] d\Phi_{nc}(p'_{nc}) \quad (19)$$

Bellman equation for the present discounted value of unemployed migrant workers:

$$rU_i = b + \lambda_{fi} \int_{p'_{0iw}}^{\infty} [W_i - U_i] d\Phi_{iw}(p'_{iw}) + \frac{m_i}{u_i} \int_{p'_{0ic}}^{\infty} [W_i - U_i] d\Phi_{ic}(p'_{ic}) \quad (20)$$

Equation (17) and (18) can be simplified and rewritten as follows:

$$rU_n = b + (\lambda_{fn} P(p'_{nw} > p'_{0nw}) + \frac{m_n}{u_n} P(p'_{nc} > p'_{0nc})) \frac{w - rU_n}{r + \delta} \quad (21)$$

$$rU_i = b + (\lambda_{fi} P(p'_{iw} > p'_{0iw}) + \frac{m_i}{u_i} P(p'_{ic} > p'_{0ic})) \frac{w - rU_i}{r + \delta} \quad (22)$$

Steady state equations for the unemployment rates of natives and migrants can be described by the following equations:

$$\dot{u}_n = \delta(f - u_n) - \lambda_{fn} P(p'_{nw} > p'_{0nw}) u_n - \frac{m_n}{u_n} P(p'_{nc} > p'_{0nc}) u_n = 0 \quad (23)$$

$$\dot{u}_i = \delta(1 - f - u_i) - \lambda_{fi} P(p'_{iw} > p'_{0iw}) u_i - \frac{m_i}{u_i} P(p'_{ic} > p'_{0ic}) u_i = 0 \quad (24)$$

At the steady state the flow into the unemployment equals to the flow out of the unemployment:

$$\frac{\delta(f - u_n)}{u_n} = \lambda_{fn} P(p'_{nw} > p'_{0nw}) + \frac{m_n}{u_n} P(p'_{nc} > p'_{0nc}) \quad (25)$$

$$\frac{\delta(1 - f - u_i)}{u_i} = \lambda_{fi} P(p'_{iw} > p'_{0iw}) + \frac{m_i}{u_i} P(p'_{ic} > p'_{0ic}) \quad (26)$$

The expression on the RHS of the equation (25) appears in the equation (21) either. The same holds for the equations (26) and (22). If we substitute the corresponding expressions in the equations (21) and (22) with the LHS of the equations (25) and (26) we can obtain the following expressions for the present discounted value of unemployed native and migrant workers:

$$U_n = \frac{b(r + \delta) + w \frac{\delta(f - u_n)}{u_n}}{r(r + \delta + \frac{\delta(f - u_n)}{u_n})} \quad (27)$$

$$U_i = \frac{b(r + \delta) + w\left(\frac{\delta(1-f-u_i)}{u_i}\right)}{r(r + \delta + \frac{\delta(1-f-u_i)}{u_i})} \quad (28)$$

Bellman equation for the present discounted value of the firm's expected profit from a filled job can be written as:

$$rE[J(p|p')] = E(p|p') - w - \delta(E[J(p|p')] - V) \quad (29)$$

Where $E[J(p|p')]$ is the present discounted value of the firm's expected profit from a filled job given the productivity signal of the worker. V is the present discounted value of the firm's expected profit from an open vacancy. The equation (28) can be rewritten so that we get the net return of the job to the firm in the LHS of the equation:

$$E[J(p|p')] - V = \frac{E(p|p') - w - rV}{r + \delta} \quad (30)$$

Under the Free Entry condition the present discounted value of the firm's expected profit from an open vacancy equals to 0:

$$\begin{aligned} rV = & -c + \frac{m_n}{v} \int_{p'_{0nc}}^{\infty} [E_p(J(p|p'_{nc})) - V] d\Phi_{nc}(p'_{nc}) + q_n \int_{p'_{0nw}}^{\infty} [E_p(J(p|p'_{nw})) - V] d\Phi_{nw}(p'_{nw}) \\ & + \frac{m_i}{v} \int_{p'_{0ic}}^{\infty} [E_p(J(p|p'_{ic})) - V] d\Phi_{ic}(p'_{ic}) + q_i \int_{p'_{0iw}}^{\infty} [E_p(J(p|p'_{iw})) - V] d\Phi_{iw}(p'_{iw}) = 0 \end{aligned} \quad (31)$$

Where the flow cost of the vacancy is c . The expected return of the vacancy to the firm consists of four parts, because there are four different groups of workers. The expected return of the vacancy to the firm is the sum of the expected returns from the groups. The expected return from a group equals to the product of the job filling rate of the group and the expected net return of the job if the job is filled by a worker of that group. Note that the firm hires a worker from a particular group only if the productivity signal of the worker is higher than the lower bound of the productivity signal for the group. Thus the firm makes sure that the net return of the job to the firm is positive when the job is filled by a worker from that group. Since the wage is determined ex-ante, the lower bounds for the groups are so that the $E(p|p'_0) - w = 0$, i.e. the RHS of the equation (28) is zero. The expected productivity of the worker equals to the predetermined wage given the worker's productivity signal equals to lower bound of the productivity signal p'_0 , i.e. $E(p|p'_0) = w$. Thus we can obtain the lower bounds of the productivity signal for the four groups by rewriting the equations (6) to (9) for the expected productivity at the lower bounds and equalizing them to the wage:

$$E(p|p'_{0nw}) = \mu + \frac{\sigma^2}{\sigma^2 + \sigma_\epsilon^2} (p'_{0nw} - \mu) = w, p'_{0nw} = (w - \mu) \frac{\sigma^2 + \sigma_\epsilon^2}{\sigma^2} + \mu \quad (32)$$

$$E(p|p'_{0nc}) = \mu + s + \frac{\sigma^2}{\sigma^2 + \sigma_\epsilon^2} (p'_{0nc} - \mu - s) = w, p'_{0nc} = (w - \mu - s) \frac{\sigma^2 + \sigma_\epsilon^2}{\sigma^2} + \mu + s \quad (33)$$

$$E(p|p'_{0iw}) = \mu - d + \frac{\sigma^2}{\sigma^2 + \sigma_\epsilon^2}(p'_{0iw} - \mu + d) = w, p'_{0iw} = (w - \mu + d)\frac{\sigma^2 + \sigma_\epsilon^2}{\sigma^2} + \mu - d \quad (34)$$

$$E(p|p'_{0ic}) = \mu - d + s + \frac{\sigma^2}{\sigma^2 + \sigma_\epsilon^2}(p'_{0ic} - \mu + d - s) = w, p'_{0ic} = (w - \mu + d - s)\frac{\sigma^2 + \sigma_\epsilon^2}{\sigma^2} + \mu - d + s \quad (35)$$

Finally, the fraction of native workers who found their job through referrals:

$$fr_n = \frac{\lambda_{nc}P(p'_{nc} > p'_{0nc})}{\lambda_{fn}P(p'_{nw} > p'_{0nw}) + \lambda_{nc}P(p'_{nc} > p'_{0nc})} \quad (36)$$

The fraction of migrant workers who found their job through referrals:

$$fr_i = \frac{\lambda_{ic}P(p'_{ic} > p'_{0ic})}{\lambda_{fi}P(p'_{iw} > p'_{0iw}) + \lambda_{ic}P(p'_{ic} > p'_{0ic})} \quad (37)$$

3.4 Wage determination

As already mentioned above, the union and the firm bargain over the wage ex-ante. The union maximizes $((f - u_N)W_N + (1 - f - u_I)W_I + u_N U_N + u_I U_I - \frac{b}{r})$, which is the sum of the present discounted values of both employed and unemployed natives and migrants. The outside option of the union is $\frac{b}{r}$ because if negotiation is not successful, everyone remains unemployed. The expression which the representative firm wants to maximize consists of four parts, because there are four different groups of employed workers. The expression is the sum of the expected returns from the four groups of employed workers. The expected return from a group equals to the product of the number of workers from that group and the expected net return of the job if the job is filled by a worker from that group. The outside option for the firm is zero.

$$\begin{aligned} & \left((f - u_n)W_n + (1 - f - u_i)W_i + u_n U_n + u_i U_i - \frac{b}{r} \right)^B \\ & \quad \left((f - u_n)fr_n \int_{p'_{0nc}}^{\infty} [E_p(J(p|p'_{nc})) - V] d\Phi_{nc}(p'_{nc}) \right. \\ & \quad + (f - u_n)(1 - fr_n) \int_{p'_{0nw}}^{\infty} [E_p(J(p|p'_{nw})) - V] d\Phi_{nw}(p'_{nw}) \\ & \quad + (1 - f - u_i)fr_i \int_{p'_{0ic}}^{\infty} [E_p(J(p|p'_{ic})) - V] d\Phi_{ic}(p'_{ic}) \\ & \quad \left. + (1 - f - u_i)(1 - fr_i) \int_{p'_{0iw}}^{\infty} [E_p(J(p|p'_{iw})) - V] d\Phi_{iw}(p'_{iw}) \right)^{1-B} \longrightarrow \max_w \quad (38) \end{aligned}$$

Where B is the bargaining power of the union. The model described above is used to numerically calculate some values of variables in the model and find an explanation of the 7.2% difference in the probabilities of using referrals when finding a job between the natives and migrants.

4 Numerical example

First, we discuss the choice of the values of the exogenous variables described in the Table 5. The mean of the workers' true productivity μ is normalized to 1. Following Shimer (2005) the total factor productivity of the matching function λ_0 is set to 1.355. Petrongolo and Pissarides

(2001) study the matching function and report that the empirical studies estimated the elasticity parameter to be from 0.5 to 0.7 when the flow of hires is used as a dependent variable, stock of unemployment and vacancies as explanatory variables. These studies assume Cobb-Douglas form of the matching function. While there are studies which use as a dependent variable not only the hires from unemployment, but the total hires and estimate η to be from 0.3 to 0.4. For the total factor productivity of the matching function η value of 0.5 is chosen as the average of the lower bound and the upper bound of the values estimated in the literature. As a fraction of the native workers we use the definition described in the data part and find the fraction of natives in the data. The values of the unemployment benefit, flow cost of the vacancy, interest rate and job destruction rate are the values used by Stupnytska (2015).

Variable	Value	Explanation
μ	1	Mean of the workers' true productivity
λ_0	1.355	Total factor productivity of the matching function
η	0.5	Elasticity with respect to the unemployment in the matching function
f	0.8783	Fraction of the native workers
h_n	90	Number of the contacts of native workers
h_i	50	Number of the contacts of migrant workers
b	0.5	Unemployment benefit
σ^2	0.1	Variance of the true productivity
σ_ϵ^2	0.1	Variance of the error term
c	0.5	Flow cost of the vacancy
γ_n	0.8783	Level of homophily between the social contacts of a native worker
γ_i	0.7	Level of homophily between the social contacts of a migrant worker
r	0.01	Interest rate
δ	0.2	Job destruction rate

Table 5: Values of the exogenous variables

Intuitively, one can expect that migrant workers have smaller social networks in the new destination country. We assume that migrant workers have smaller number of the contacts than natives. 90 and 50 are the numbers of the contacts of worker with high and low social capital in the study by Stupnytska (2015). The variance of the true productivity and the variance of the error term are chosen so that most of the observations have positive productivity. Titzmann and Silbereisen (2009) study the friendship homophily among the emigrant adolescents from Soviet union to Germany and find high levels of friendship homophily. Levels of friendship homophily among the newcomers was 75% and 65% among experienced. For the level of homophily between the social contacts of a migrant worker the average of these two is used. For the level of homophily between the social contacts of a migrant worker we use the fraction of natives assuming natives form contacts randomly.

The model is calibrated using the (23) and (24) steady state equations for the unemployment rates of natives and migrants, Free Entry condition described in the equation (31), and the fractions of migrant and native workers who found their job through referrals from the equations (37) and (36). The unemployment rates of natives and migrants are calculated according to the ILO guidelines from the SOEP data of the years 2002 to 2008. As a fraction of natives/migrants who found job through referrals the predicted probabilities of the average native*/migrant* are used. The calibration results are presented in the Table 6. Further, Table 7 shows some of the

Variable	Value	Explanation
s	0.2018	Average productivity difference between the workers with and without a referral
d	0.2150	Average productivity difference between the native and migrant workers
a	0.0787	Vacancy arrival rate
v	1.2993	Vacancy rate
B	0.7821	Bargaining power of the union
u_n/f	0.0736	Unemployment rate of native workers
$u_i/1-f$	0.1364	Unemployment rate of migrant workers
fr_n	0.2719	Fraction of natives who found job through referrals
fr_i	0.3440	Fraction of migrants who found job through referrals

Table 6: Calibration results

values of the endogenous variables.

Variable	Value	Explanation
λ_{fn}	6.07340	Job finding rate in the formal market for natives
λ_{fi}	11.98538	Job finding rate in the formal market for migrants
λ_{nc}	1.05383	Job finding rate through referrals for natives
λ_{ic}	1.54712	Job finding rate through referrals for migrants
w	1.116199367	Wage
q_n	0.30231	Job filling rate in the formal market for natives
q_i	0.15319	Job filling rate in the formal market for migrants
m_n	0.06815	Number of matches through referrals for natives
m_i	0.02569	Number of matches through referrals for migrants

Table 7: Values of the endogenous variables

Figure 1 depicts the probability density functions and the lower bounds of the productivity signals for the four groups. The blue dash line is the PDF of the productivity signal for the native workers matched through referrals. The vertical blue dash line is the lower bound of the productivity signal for the native workers matched through referrals. The surface area bounded with these two blue dash lines is the probability that the native workers have productivity signal higher than the lower bound when matched with the firm through referrals. The blue line is the PDF of the productivity signal for the native workers matched through formal channels. The vertical blue line is the lower bound of the productivity signal for the native workers matched through formal channels. The surface area bounded with these two blue lines is the probability that the native workers have productivity signal higher than the lower bound when matched with the firm through formal channels. Similarly, The green dash line is the PDF of the productivity signal for the migrant workers matched through referrals. The vertical green dash line is the lower bound of the productivity signal for the migrant workers matched through referrals. The surface area bounded with these two green dash lines is the probability that the migrant workers have productivity signal higher than the lower bound when matched with the firm through referrals. The green line is the PDF of the productivity signal for the migrant workers matched through formal channels. The vertical green line is the lower bound of the productivity signal for the migrant workers matched through formal channels. The surface area bounded with these two green lines is the probability that the migrant workers have productivity signal higher than the lower bound when matched with the firm through formal channels. The corresponding

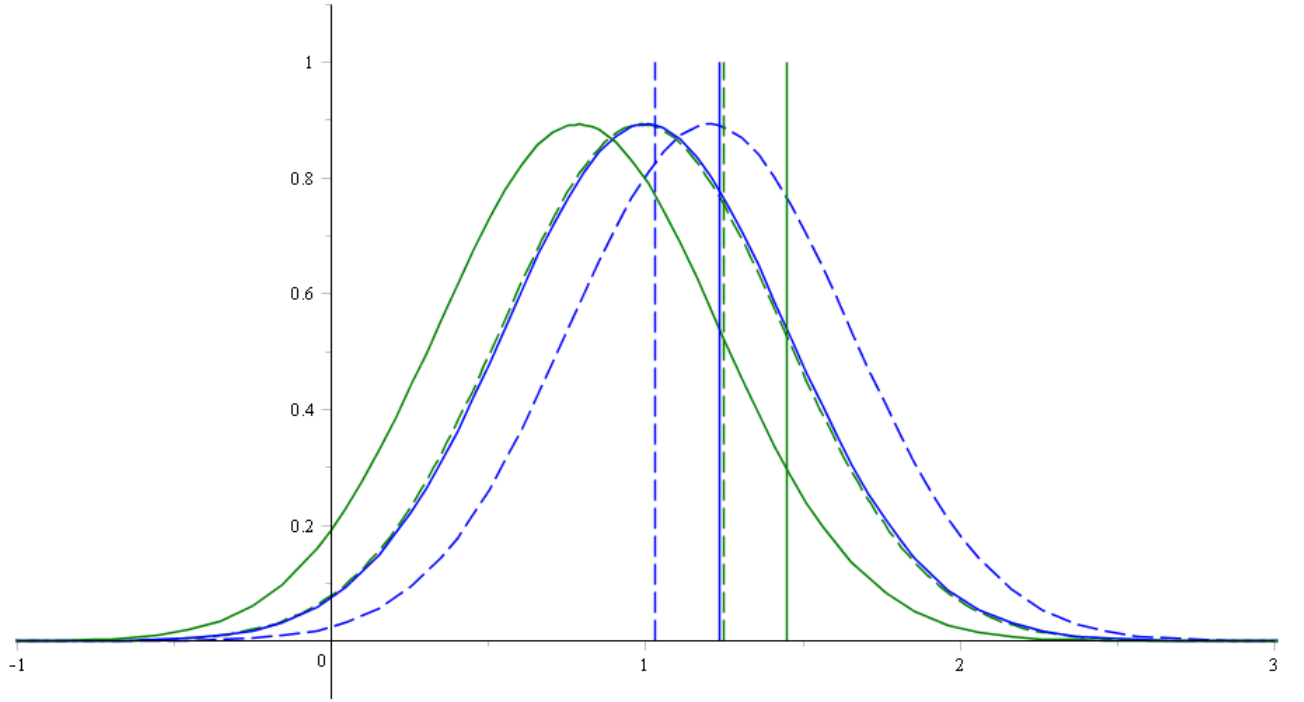


Figure 1: Probability density functions and the lower bounds of the productivity signals.

probabilities and the values of the lower bounds of the four groups are reported in the Table 8.

p'_{0nw}	$P(p'_{nw} > p'_{0nw})$	p'_{0nc}	$P(p'_{nc} > p'_{0nc})$	p'_{0iw}	$P(p'_{iw} > p'_{0iw})$	p'_{0ic}	$P(p'_{ic} > p'_{0ic})$
1.2324	0.3016	1.0305	0.6491	1.4474	0.0693	1.2455	0.2815

Table 8: Values of lower bounds and probabilities to be hired after the match

Table 8 shows that the probability that the native workers are hired after they are matched with the firm through referrals is 0.6491, which is approximately 2.15 times more than the probability for the native workers matched through formal channels. The probability for the native workers matched through formal channels is 0.3016. The probability that the migrant workers are hired after they are matched with the firm through referrals is 0.2815, 4.06 times more than the probability for the migrant workers matched through formal channels, which equals to 0.0693. So the gain from finding job through referrals for the natives is lower than the gain for the migrants, which can be the reason of the 7.2% difference in the probabilities of using referrals when finding a job between the natives and migrants.

5 Conclusions

Empirical analysis of the SOEP data from 2002 to 2008 show that 41.21% of migrant workers found their current job through referrals. While 31.79% of native workers found their current job through referrals. Estimation results of the panel probit model with random effects show that a part of the 9.42% can be explained by the control variables including characteristics of the individuals and firms. But still there is 7.2% statistically significant difference of the predicted probabilities of using referrals between average migrant* and average native* workers. Intuitively,

one can expect that migrant workers have smaller social networks in the new destination country and by that smaller probability of finding a job through referrals, but our empirical observation shows that migrant workers are more likely to find a job through referrals even after introducing controls. In order to explain this puzzle, this paper presents a search and matching model with heterogeneous worker groups and several search channels. Even though in the model native workers have more social contacts than migrants, migrant workers are still more likely to find job through referrals. Results from the calibration of the model show that as expected, average productivity difference between the native and migrant workers is positive and equals to 0.2018. Moreover, average productivity difference between the workers with and without referrals is also positive and equals to 0.2150. In this setting, the firm does not observe the real productivity of the worker, the firm observes the productivity signal of the worker, the group and matching channel of the worker. The worker is hired if the expected productivity of the worker given the productivity signal is higher than the lower bound of the productivity signal for the group of the worker. The lower bounds of the groups are determined based on the ex-ante bargained wage. Thus, given the distributions of the productivity signals for the four groups, we obtain the probability of being hired after the match for the four groups.

The results presented in the Table 8 show that the probability that the native workers are hired after they are matched with the firm through referrals is approximately 2.15 times more than the probability for the native workers matched through formal channels. The probability that the migrant workers are hired after they are matched with the firm through referrals is 4.06 times more than the probability for the migrant workers matched through formal channels. So the gain from finding job through referrals for the migrants is higher than the gain for the natives. In conclusion, we propose the following explanation to the reason of the different frequency of finding job through referrals between natives and migrants. Migrant workers have low chances of being hired therefore they gain more from being matched to a job vacancy through referrals. Native workers have good chances of being hired even if they are matched through formal channels therefore they gain less from referrals.

References

- [1] Andersen, E.B., 1970. Asymptotic properties of conditional maximum-likelihood estimators.
- [2] Andersen, E.B., 1971. The asymptotic distribution of conditional likelihood ratio tests. *Journal of the American Statistical Association*, 66(335), pp.630-633.
- [3] Bentolila, S., Michelacci, C. and Suarez, J., 2010. Social contacts and occupational choice. *Economica*, 77(305), pp.20-45.
- [4] Brown, M. and Setren, E., G. Topa (2014), "Do informal referrals lead to better matches? Evidence from a firm's employee referral system," Unpublished manuscript, Federal Reserve Bank of New York.
- [5] Butler, J.S. and Moffitt, R., 1982. A computationally efficient quadrature procedure for the one-factor multinomial probit model. *Econometrica: Journal of the Econometric Society*, pp.761-764.
- [6] Cahuc, P. and Fontaine, F., 2009. On the efficiency of job search with social networks. *Journal of Public Economic Theory*, 11(3), pp.411-439.
- [7] Calvo-Armengol, A. and Jackson, M.O., 2004. The effects of social networks on employment and inequality. *The American Economic Review*, 94(3), pp.426-454.
- [8] Cappellari, L. and Tatsiramos, K., 2013. With a little help from my friends? quality of social networks, job finding rates and job match quality. Working Paper.
- [9] Chamberlain, G., 1980. Analysis of Covariance with Qualitative Data. *The Review of Economic Studies* 47(1), pp. 225-238.
- [10] DeGroot, M.H., 2005. *Optimal statistical decisions* (Vol. 82). John Wiley Sons.
- [11] Drever, A.I. and Hoffmeister, O., 2008. Immigrants and Social Networks in a Job Scarce Environment: The Case of Germany. *International Migration Review*, 42(2), pp.425-448.
- [12] Dustmann, C., Glitz, A. and Schönberg, U., 2011. Referral-based job search networks.
- [13] Eisnecker, P. and Schacht, D., 2016. Half of the refugees in Germany found their first job through social contacts. *DIW Economic Bulletin*, 6(34/35), pp.414-421.
- [14] Galenianos, M., 2013. Learning about match quality and the use of referrals. *Review of Economic Dynamics*, 16(4), pp.668-690.
- [15] Ganzeboom, H.B.G. De Graaf, P.M., Treiman, D.J., 1992. A Standard International Socio-Economic Index of Occupational Status. *Social Science Research* 21 (1), 1-56.
- [16] Greene, W.H., 2007. *Econometric analysis*. 6th. ed. Upper Saddle River, NJ: Pearson/Prentice Hall.

- [17] Helberger, Christof 1988: Eine Überprüfung der Linearitätsannahme der Humankapitaltheorie. In: H.-J. Bodenhöfer (ed.): Bildung, Beruf, Arbeitsmarkt, pp. 151-170.
- [18] Johnson, N.L. and Kotz, S., 1972. Distributions in statistics. Continuous univariate distributions, 2.
- [19] Jovanovic, B., 1979. Job matching and the theory of turnover. *The Journal of Political Economy*, pp.972-990.
- [20] Jovanovic, B., 1984. Matching, turnover, and unemployment. *The Journal of Political Economy*, pp.108-122.
- [21] Kaas, L. and Manger, C., 2012. Ethnic discrimination in Germany's labour market: a field experiment. *German Economic Review*, 13(1), pp.1-20.
- [22] Kugler, A.D., 2003. Employee referrals and efficiency wages. *Labour economics*, 10(5), pp.531-556.
- [23] Lancaster, T., 2000. The incidental parameter problem since 1948. *Journal of econometrics*, 95(2), pp.391-413.
- [24] Lancee, B., 2016. Job search methods and immigrant earnings: A longitudinal analysis of the role of bridging social capital. *Ethnicities*, 16(3), pp.349-367.
- [25] Maddala, G.S., 1987. Limited Dependent Variable Models Using Panel Data. *Journal of Human Resources* 22(3), pp.307-338.
- [26] Marsden, P.V., 1988. Homogeneity in confiding relations. *Social networks*, 10(1), pp.57-76.
- [27] Montgomery, J.D., 1991. Social networks and persistent inequality in the labor market. Center for Urban Affairs and Policy Research.
- [28] Naylor, J.C. and Smith, A.F., 1982. Applications of a method for the efficient computation of posterior distributions. *Applied Statistics*, pp.214-225.
- [29] Neyman, J. and Scott, E.L., 1948. Consistent estimates based on partially consistent observations. *Econometrica: Journal of the Econometric Society*, pp.1-32.
- [30] Petrongolo, B. and Pissarides, C.A., 2001. Looking into the black box: A survey of the matching function. *Journal of Economic literature*, 39(2), pp.390-431.
- [31] Phelps, E.S., 1972. The statistical theory of racism and sexism. *The American Economic Review*, 62(4), pp.659-661.
- [32] Pistaferri, L., 1999. Informal networks in the Italian labor market. *Giornale degli economisti e annali di economia*, pp.355-375. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp.283-301.
- [33] Pissarides, C.A., 2000. *Equilibrium unemployment theory*. MIT press.

- [34] Rasch, G., 1960. Probabilistic models for some intelligence and achievement tests. Copenhagen: Danish Institute for Educational Research.
- [35] Rees, A., 1966. Information networks in labor markets. *The American Economic Review*, 56(1/2), pp.559-566.
- [36] Schwarze, Johannes 1991: Ausbildung und Einkommen von Männern - Einkommensfunktionsschätzungen für die ehemalige DDR und die Bundesrepublik Deutschland. In: *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, (24), pp. 63-69.
- [37] Shimer, R., 2005. The cyclical behavior of equilibrium unemployment and vacancies. *The American Economic Review*, 95(1), pp.25-49.
- [38] StataCorp. 2013, *Stata 13 Base Reference Manual*. College Station, TX: Stata Press.
- [39] Stupnytska, Y., 2015. Asymmetric information in a search model with social contacts.
- [40] Stupnytska, Y. and Zaharieva, A., 2015. Explaining U-shape of the referral hiring pattern in a search model with heterogeneous workers. *Journal of Economic Behavior & Organization*, 119, pp.211-233.
- [41] Titzmann, P.F. and Silbereisen, R.K., 2009. Friendship homophily among ethnic German immigrants: A longitudinal comparison between recent and more experienced immigrant adolescents. *Journal of Family Psychology*, 23(3), p.301.