Social Contacts and Referrals in Search and Matching Models of the Labour Market

A dissertation submitted to the Faculty of Business Administration and Economics of Bielefeld University in partial fulfillment of the requirements for the doctoral degree in Economics

submitted by
Sevak Alaverdyan

FIRST SUPERVISOR: J.-Prof. Dr. Anna Zaharieva
SECOND SUPERVISOR: Prof. Dr. Herbert Dawid

March 2020

Acknowledgements

The writing of the dissertation would have been impossible without the support of many people. First of all, I want to thank my first supervisor J.-Prof. Dr. Anna Zaharieva for her enormous support. She provided continuous guidance over my doctoral studies. I am particularly grateful to her for the knowledge and skills, time spent in discussions, motivation and inspiration to research, and the great moral support. I want to thank my second supervisor Prof. Dr. Herbert Dawid for his guidance, comments, and suggestions related to research. Moreover, I am especially grateful to him for the great support to overcome all the possible administrative obstacles which arose during my doctoral studies. I also thank J.-Prof. Dr. Tim Hellmann for his valuable feedback during the seminar dealing with social networks. Special thanks to Prof. Dr. Gerald Willmann for the useful feedback on my research and willingness to support whenever needed.

Furthermore, I would like to thank the faculty members of BiGSEM and IMW, and all the Ph.D. candidates at BiGSEM for their comments and suggestions during various discussions and my presentations at the colloquium. Special thanks to Lasha Chochua and Quji Bichia for accommodating, and providing guidance and support from the very first day of my studies. Next, I would like to thank Erdenebulgan Damdinsuren and Ashenafi Guta for their friendship, great help through discussions, and the productive and positive atmosphere in the office we shared together.

I want to thank German Academic Exchange Service (DAAD) for funding my doctoral studies within the Graduate School Scholarship Program (GSSP), opportunity and scholarship provided to participate in the 6th Lindau Meeting on Economic Sciences, where I had an opportunity to meet and follow lectures and discussions by 20 Nobel Laureates in Economics.

I am very grateful to Dr. Michael Stops and Assistant Prof. Dr. Michele Battisti for their insightful comments and suggestions during our meetings and discussions on various occasions. Also, I thank the participants of conferences, workshops and seminars Lyon, Bielefeld, Cambridge, Alicante, Trier, Cosenza, Lindau, Dortmund, Nuremberg, Yerevan, and Dilijan. Thanks to the funding provided by the Central Bank of Armenia (CBA) I had an opportunity to present my research to the research team of CBA and get valuable comments and feedback.

I would like to express my gratitude to my lecturers and classmates from the International School of Economics at TSU (ISET) for the knowledge, and constant support during my research activities. Moreover, I am grateful to ISET for providing financial assistance at the stage of applying to Ph.D. programs. The assistance enabled me to successfully get admitted to BiGSEM.

Last but not least, I am thankful to my parents Harutyun and Nelli, my sister Tatevik, and my friends for their encouragement and continuous support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This dissertation consists of three studies devoted to the role of social contacts and referrals in search and matching models within the labour market. Social networks serve as an important mechanism for the transmission of information between employers and employment seekers. Empirical evidence shows that 30-50% of employees find jobs through social contacts, [1] therefore, social networks play an important role in mitigating informational problems common to the labour market. Referrals provide employers with information about the inconspicuous characteristics of employment seekers (see, Montgomery (1991), Galenianos (2013), Dustmann et al. (2016); social contacts of job seekers provide them with information about vacancies, thus reducing search frictions(Calvó-Armengol and Zenou (2005), Horvath (2014)); and referrals enable employers to elicit effort less costly, since they provide screening and monitoring of new hires through their social contacts(Kugler (2003)). Various theoretical models are developed and empirical studies are conducted to explore the effect of referral hiring on labour market outcomes such as wages, probability of being employed, productivity of the workers and the firms, etc. The main goal of this thesis is to investigate the effect of referrals on labour market outcomes, with a special focus on the use of referrals by immigrant workers.

To evaluate the effect of referrals on labour market outcomes it is essential to understand the structure of social networks, the frequency and form of their usage by different demographic groups.[2] Whereas most studies document the positive effect of referrals on the probability of jobseekers successfully finding employment; findings regarding wages are not as conclusive. While Montgomery (1991), Simon and Warner (1992), Kugler (2003), Dustmann et al. (2016) emphasize the positive effect referrals have on wages, Pistaferri (1999), Addison and Portugal (2002), Bentolila et al. (2010) and Zaharieva (2018) argue otherwise. In particular, Bentolila et al. (2010) and Zaharieva (2018) find that the use of referrals can be associated with occupational mismatch. Several other studies find both positive and negative effect on wages within the same study. [3]

[1] see, for example, Holzer (1988), Granovetter (1995), Pistaferri (1999), Kugler (2003), Pellizzari (2010), Drever and Hoffmeister (2008), Bentolila et al. (2010).

[2] see detailed literature review in Ioannides and Datcher Loury (2004).

[3] For example, Sylos Labini (2005), Pellizzari (2010) and Tumen (2016).

Throughout this thesis we find both positive (Chapter 2) and negative (Chapter 3) effects of referrals on match quality. The findings in Chapter 2 suggest that both native and immigrant workers have a higher probability of being employed if they are matched with firms by referral rather than by formal search channels. Moreover, recommended workers have higher productivity rates once employed. It is also possible to say that immigrants rely more on referrals than natives do due to the probability of them being hired via a referral being higher than through other channels. The findings in Chapter 3 reveal that employees, both native and immigrant, are more likely to be mismatched if hired through a referral than their counterparts, who were hired via formal channels.

While the studies mentioned above observe the effect of using referrals on labour market outcomes, Galeotti and Merlino (2003) explore the effect of labour market conditions on the use of networks and their efficiency in matching employees to vacant positions. Chapter 4 is closely related to the study by Galeotti and Merlino (2003) however, it aims to find the effect of labour market conditions on the frequency of using referrals. In Chapter 4 we explore the effects of change in productivity on search intensities of workers and firms; hence, contributing to the literature on labour market with endogenous search effort by Diamond (1982), Mortensen and Pissarides (1994), Mortensen and Pissarides (1999), and Pissarides (2000).

Chapter 2 investigates both empirically and theoretically the reasons contributing to the the difference in the frequency of referral usage between natives and immigrants in German labour market. It is assumed that migrant workers have smaller social networks in their new country destination and therefore a smaller probability of finding employment through referrals. However, empirical analysis of the German Socio-Economic Panel (SOEP) data from 2002 to 2008 shows that 41.21% of migrant workers and 31.79% of native workers found their current employment through referrals. Estimation results show that the difference of 9.42% between the two can be partly explained by the characteristics of different individuals and firms. In addition, we find that referral hiring is more frequent for younger, less educated workers and it is more widespread in smaller firms. These findings are inline with the results of Drever and Hoffmeister (2008). Furthermore, estimation results of the panel probit model with random effects show a 7.26% statistically significant difference of the predicted probabilities of using referrals between migrant and native workers, which cannot be explained by the characteristics of the individuals and firms.

In order to explain this puzzle, this paper presents a search and matching model with heterogeneous worker groups and several search channels. There are two groups of workers: natives and immigrants, who are matched to the vacancy through referrals or formal search channels. In the model, immigrants have fewer contacts than natives. Similar to Montgomery (1991), social networks are characterised by the inbreeding homophily bias based on migration background. We assume that compared to the workers without referral, recommended workers produce different signal to the firm. The firm observes the noisy signal of the productivity, the nationality and the search channel of the worker to form unbiased expectations about the true productivity of the worker. The probabilities to be hired for the two worker groups and different search channels are determined using ex-ante union bargained wage and expectations of the firm. 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.

Calibration results of the model show that even when migrant workers have a smaller sized social network, they gain more from recommendations because their employment chances are initially much lower than the employment chances of native workers. The probability that a native worker is hired after they are matched with the firm through referrals is approximately 1.74 times higher than the probability them being hired through formal channels. The probability that of a migrant worker being hired after they are matched with the firm through referrals is 3.29 times greater than the probability of a migrant worker who is matched through formal channels.

In the study presented in Chapter 3, we examine the link between search channels that workers use to find employment and the probability of occupational mismatch in their new position. Our specific focus is on the differences between native and immigrant workers and we use data from the SOEP over the period between 2000-2014. First, we document that referral hiring via social networks is the most frequent single channel of job generation in Germany; relatively speaking, referrals are used more frequently by immigrant workers compared to natives. Secondly, our data reveals that referral hiring is associated with the highest rate of occupational mismatch among all channels in Germany. This result is inline with the findings highlighted by Bentolila et al. (2010) for the United States.

We combine these findings and use them to develop a theoretical search and matching model with two ethnic groups of workers (natives and immigrants), two search channels (formal and referral hiring) and two occupation types. When modelling social networks, we take into account ethnic and professional homophily in the formation of links. Our model predicts that immigrant workers face a stronger risk of unemployment and often rely on recommendations from their friends and relatives as a last resort. Furthermore, higher rates of referral hiring produce more frequent occupational mismatch of the immigrant population compared to the native population.

We test this prediction empirically by performing a Blinder-Oaxaca decomposition and confirm that intensive network hiring contributes significantly to higher rates of occupational mismatch among immigrants, but the contribution is quantitatively small. The differences in the search strategies explain about 1% of the 15.5% gap in the mismatch rates. Finally, we document that the gaps in the incidences of referrals as well as mismatch rates are reduced among second generation immigrants, indicating some degree of integration in the German labour market.

Chapter 4 of the dissertation explores the effects of productivity change on the relative frequency of using referrals by firms and workers. On the one hand, it is relatively easier for workers to find jobs during expansions and therefore they rely less on their social networks, resulting in a lower frequency of hiring through referrals during expansions. On the other hand, firms have difficulties in filling their open positions during expansion periods and therefore , they rely more heavily on their social networks, resulting in the referral hiring system becoming more dominating. Empirical analysis of the IAB Job Vacancy Survey data between the years 2000 to 2013 and SOEP data from 2000 to 2014, show that in the long run there is a positive correlation between the GDP and the proportion of workers hired through their social networks.

In order to explain the effect of the productivity change on the search and matching strategies of firms and workers, this paper presents a search and matching model with several search channels. According to Pissarides (2000), firms choose to concentrate their advertising efforts through formal search channels in order to maximize the asset value of an open vacancy. Employment seekers choose to intensify their search efforts through formal channels in an attempt to maximize the asset value of being unemployed. Galeotti and Merlino (2003) assume that workers invest in their networks in order to keep their networks active. Furthermore, incentives in networking are related to labour market conditions, in particular, to the separation rate. This model uses the more conventional approach where the flow of information about vacancies through social networks is exogenous.

Calibration results of the model show that even though productivity increases during expansions, formal advertisement is not profitable for the firms because of higher competition rates, as well as higher wage expectations. As a result, the firms advertise less and therefore the proportion of referral hiring increases. When productivity increases, employment becomes more gainful, encouraging workers to exert more effort in searching for employment through formal channels. As a result, the proportion of referral hiring should decrease. Both the estimation and the calibration results indicate that the firm-side effect dominates. Similar to Pissarides (2000), our findings suggest that the search intensity of workers through formal channels is procyclical in the presence of referral hiring.

Contributions

Chapter 3 of this dissertation is a joint work with J.-Prof. Dr. Anna Zaharieva. During this research project, J.-Prof. Dr. Zaharieva was my first supervisor. My major contribution was running empirical estimation based on the data. While J.-Prof. Dr. Zaharieva developed the theoretical model and conducted numerical calibrations. It is impossible to attribute the rest of the work within the project to any particular person. It is an outcome of many discussions and our joint efforts.

Declaration

I hereby declare that I am aware of the current doctoral regulations of the faculty. The thesis is my and my co-author's original work. Any other contributions are marked as such. All other sources and utilized tools are listed. No third parties benefited financially (in a direct or indirect way) from work related to the content of this thesis. The thesis was never submitted as part of an examination before.

Chapter 2

Why do migrant workers rely more often on referrals?

2.1 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 (see Granovetter (1995)). 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.

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 two groups of workers:natives and immigrants. The workers are matched to 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 (1972) 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. (2016) 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 1.74 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 3.29 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.

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. (2016), Galenianos (2013)), while others find the effect to be negative (Pistaferri (1999), Bentolila et al. (2010)).

Social contacts tend to occur among workers with similar characteristics. Moreover, 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. (2016) build their model on the Jovanovic (1979), and Jovanovic (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. (2016) 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 et al. (2016) 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%. Dawid and Gemkow (2014) employ a closed agent-based macroeconomic simulation model to study how social networks contribute to wage inequality. They find that due to the referral hiring workers with high specific skills are matched to the high productive firms. Therefore, the workers who find their job through referrals get on average higher wages than those who find job through other channels.

Cappellari and Tatsiramos (2015) 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 (2015) 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.

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.

2.2 Empirical Analysis

2.2.1 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.

2.2.2 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 2.2, 32.9% of observed respondents replied, that they found their current job through referrals. On the one hand, according to the Table 2.1, there is significant 9.42% difference between the native and migrant individuals which found the current job through referrals . On the other 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.

Table 2.2 presents descriptive statistics of the variables used in the estimations. Besides the statistics about the overall sample observed, Table 2.2 includes descriptive statistics separately for the natives and the migrants, to better understand the differences between

[1] for detailed description see Helberger (1988) and Schwarze et al. (1991).

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

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

Table 2.1: Percentage of individuals found the current job through referrals by citizenship\migration status.

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
$ISEI_{it}$	44.83	38.38	44.04

Table 2.2: 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, 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 and they are employed at smaller-size firms.

2.2.3 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 et al. (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), Andersen (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]

2.2.4 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 2.3 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 6.3% 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 43.4% for the firm size category "less than 5" to 27.5% for the firm size category "2000 and more." For a native with average characteristics the corresponding probabilities are 36.5% and 21.9% respectively.

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

[3] See details at the Stata 13 Base Reference Manual. StataCorp (2013)

[4] Furthermore, 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.

Variable	Pooled probit Coefficient	Pooled probit Marginal effects	Panel probit RE Coefficient	Panel probit RE Marginal effects
Const	0.741***		0.802***	
	(0.155)		(0.168)	
EDU_{it}	-0.019***	-0.0066***	-0.020***	-0.0069***
	(0.007)	(0.0024)	(0.007)	(0.0025)
AGE_{it}	-0.026***	-0.0014**	-0.028***	-0.0015*
	(0.009)	(0.0006)	(0.010)	(0.0006)
AGE_{it}^2	0.00032**		0.00035**	
	(0.00012)		(0.00013)	
MIG_{it}	0.172***	0.062***	0.178***	0.063***
	(0.043)	(0.016)	(0.046)	(0.017)
$FSIZE_{it}$				
5 to 19	-0.128***	-0.048***	-0.141***	-0.052***
	(0.048)	(0.018)	(0.051)	(0.019)
20 to 99	-0.223***	-0.081***	-0.243***	-0.088***
	(0.049)	(0.018)	(0.053)	(0.019)
100 to 199	-0.236***	-0.086***	-0.252***	-0.091***
	(0.060)	(0.022)	(0.065)	(0.023)
200 to 1999	-0.335***	-0.120***	-0.364***	-0.127***
	(0.053)	(0.019)	(0.057)	(0.020)
2000 and more	-0.396***	-0.139***	-0.430***	-0.148***
	(0.054)	(0.019)	(0.058)	(0.020)
$TOJCH_{it}$				
job after break	-0.062	-0.022	-0.059	-0.020
	(0.052)	(0.019)	(0.056)	(0.020)
job with new	0.118**	0.043**	0.136**	0.049**
employer	(0.050)	(0.018)	(0.054)	(0.019)
company taken	-0.732***	-0.207***	-0.793***	-0.210***
over	(0.110)	(0.025)	(0.121)	(0.026)
changed job	-0.904***	-0.237***	-0.977***	-0.238***
same firm	(0.083)	(0.021)	(0.092)	(0.021)
$ISEI_{it}$	-0.0052***	-0.0018***	-0.0058***	-0.0020***
	(0.0011)	(0.00038)	(0.0012)	(0.00039)
Observations	9670	9670	9670	9670
$McFadden's\ \ R^2$	0.0574		0.0599	

Significance level: * if $p < 0.10$, ** if $p < 0.05$, *** if $p < 0.01$.
Standard errors are in parantheses.

Table 2.3: Estimated coefficients and marginal effects.

referrals of an individual with average characteristics is 30.7% 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 28.7%. But in case of job with new employer the probability is significantly different and equals to 35.7%. 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 9.7%, in the latter case it is 6.9%. And finally, an individual with average characteristics is less likely to find job through referrals if the individual's $ISEI_{it}$ is higher.

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 2.4 shows different predicted probabilities using pooled probit model and panel probit RE.

	Predicted probabilities	
	Pooled probit	Panel probit RE
Migrants at means	0.354	0.339
Natives at means	0.292	0.276
Average migrant	0.393	0.380
Average native	0.288	0.271
Average of migrants	0.413	0.403
Average of natives	0.318	0.306
All migrants	0.382	0.371
All natives	0.322	0.310
Average migrant*	0.363	0.348
Average native*	0.291	0.275

Table 2.4: Predicted probabilities

"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 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 7.26% difference between "average migrant*" and "average native*."

2.3 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

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

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.

2.3.1 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.

$$p \sim N(\mu, \sigma^2), \epsilon \sim N(0, \sigma_\epsilon^2) \tag{2.1}$$

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 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:

$$p_n \sim N(\mu, \sigma^2), p_i \sim N(\mu - d, \sigma^2) \tag{2.2}$$

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:

$$p'_{nw} = p + \epsilon, p'_{nw} \sim N(\mu, \sigma^2 + \sigma_\epsilon^2) \tag{2.3}$$

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

$$p'_{nc} = p + s + \epsilon, p'_{nc} \sim N(\mu + s, \sigma^2 + \sigma_\epsilon^2) \tag{2.4}$$

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

$$p'_{iw} = p - d + \epsilon, p'_{iw} \sim N(\mu - d, \sigma^2 + \sigma_\epsilon^2) \tag{2.5}$$

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

$$p'_{ic} = p - d + s + \epsilon, p'_{ic} \sim N(\mu - d + s, \sigma^2 + \sigma_\epsilon^2) \tag{2.6}$$

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:

$$E(p|p'_{nw}) = \mu + \frac{\sigma^2}{\sigma^2 + \sigma_\epsilon^2}(p'_{nw} - \mu) \tag{2.7}$$

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

$$E(p|p'_{nc}) = \mu + s + \frac{\sigma^2}{\sigma^2 + \sigma_\epsilon^2}(p'_{nc} - \mu - s) \tag{2.8}$$

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

$$E(p|p'_{iw}) = \mu - d + \frac{\sigma^2}{\sigma^2 + \sigma_\epsilon^2}(p'_{iw} - \mu + d) \tag{2.9}$$

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

$$E(p|p'_{nc}) = \mu - d + s + \frac{\sigma^2}{\sigma^2 + \sigma_\epsilon^2}(p'_{nc} - \mu + d - s) \tag{2.10}$$

2.3.2 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 $m_{fj}(u_j, v)$ have Cobb-Douglas form in case of the matching through formal channels.

$$m_{fn}(u_n, v) = \lambda_0 u_n^\eta v^{1-\eta}, m_{fi}(u_i, v) = \lambda_0 u_i^\eta v^{1-\eta} \tag{2.11}$$

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 repectively:

$$\lambda_{fn} = \frac{\lambda_0 v^{(1-\eta)}}{u_n^{(1-\eta)}}, \lambda_{fi} = \frac{\lambda_0 v^{(1-\eta)}}{u_i^{(1-\eta)}} \tag{2.12}$$

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:

$$q_n = \frac{\lambda_0 u_n^\eta}{v^\eta}, q_i = \frac{\lambda_0 u_i^\eta}{v^\eta} \tag{2.13}$$

Next, we explain the mechanism of referral hiring. According to Montgomery (1991) social networks are often characterised by the inbreeding bias (homophily). This means that social links are often formed between similar workers. Following this idea we expect that native workers will have a large fraction of other native workers in their social network. Let this proportion be denoted by γ_n. This also means that they have $(1-\gamma_n)$ contacts with immigrant workers. Similar, let γ_i denote the fraction of immigrant workers in the social network of immigrant workers. Then the fraction of native workers in the social network of immigrant workers is $(1-\gamma_i)$.

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

$$\begin{aligned} m_n = va((f-u_n)[1-(1-\frac{u_n}{f})^{\gamma_n h_n}(1-\frac{u_i}{1-f})^{(1-\gamma_n)h_n}]\frac{\gamma_n \frac{u_n}{f}}{\gamma_n \frac{u_n}{f} + (1-\gamma_n)\frac{u_i}{1-f}} \\ + (1-f-u_i)[1-(1-\frac{u_i}{1-f})^{\gamma_i h_i}(1-\frac{u_n}{f})^{(1-\gamma_i)h_i}]\frac{(1-\gamma_i)\frac{u_n}{f}}{\gamma_i \frac{u_i}{1-f} + (1-\gamma_i)\frac{u_n}{f}} \end{aligned} \tag{2.14}$$

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 $\dfrac{\gamma_n \frac{u_n}{f}}{\gamma_n \frac{u_n}{f} + (1-\gamma_n)\frac{u_i}{1-f}}$. 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 $\dfrac{(1-\gamma_n)\frac{u_i}{(1-f)}}{\gamma_n \frac{u_n}{f} + (1-\gamma_n)\frac{u_i}{1-f}}$.

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:

$$\begin{aligned} m_i = va((1-f-u_i)[1-(1-\frac{u_i}{1-f})^{\gamma_i h_i}(1-\frac{u_n}{f})^{(1-\gamma_i)h_i}]\frac{\gamma_i \frac{u_i}{1-f}}{\gamma_i \frac{u_i}{1-f} + (1-\gamma_i)\frac{u_n}{f}} \\ + (f-u_n)[1-(1-\frac{u_n}{f})^{\gamma_n h_n}(1-\frac{u_i}{1-f})^{(1-\gamma_n)h_n}]\frac{(1-\gamma_n)\frac{u_i}{(1-f)}}{\gamma_n \frac{u_n}{f} + (1-\gamma_n)\frac{u_i}{1-f}} \end{aligned} \tag{2.15}$$

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

$$\lambda_{nc} = \frac{m_n}{u_n}, \lambda_{ic} = \frac{m_i}{u_i} \tag{2.16}$$

2.3.3 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:

$$rW_n = w - \delta(W_n - U_n) \quad \text{or} \quad W_n - U_n = \frac{w - rU_n}{(r+\delta)} \tag{2.17}$$

Where δ is the exogenous job destruction rate, and U_n is the present discounted value of unemployed native workers. So all employed workers earn a wage w; at the exogenous rate δ they lose their job and become unemployed. Hence the expected capital loss from losing a job for a native worker is $\delta(W_n - U_n)$.

Bellman equation for the present discounted value of an employed migrant worker W_i

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

can be written as:

$$rW_i = w - \delta(W_i - U_i) \quad \text{or} \quad W_i - U_i = \frac{w - rU_i}{(r + \delta)} \tag{2.18}$$

Where U_i is the present discounted value of unemployed migrant workers. Similar to native workers, migrant workers also earn a wage w and at the exogenous rate δ they lose their job and become unemployed. But the expected capital loss from losing a job for a migrant worker is $\delta(W_i - U_i)$.

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

$$rU_n = b + \lambda_{fn} \int_{p'_{0nw}}^{\infty} [W_n - U_n] d\Phi_{nw}(p'_{nw}) + \lambda_{nc} \int_{p'_{0nc}}^{\infty} [W_n - U_n] d\Phi_{nc}(p'_{nc}) \tag{2.19}$$

The unemployed native worker gets unemployment benefit b, and expects to move into employment through formal channels at job finding rate λ_{fn}. The unemployed native worker also expects to move into employment through referrals at rate λ_{nc}. The expected capital gain of the unemployed native worker from finding a job through formal channels is $\lambda_{fn} \int_{p'_{0nw}}^{\infty} [W_n - U_n] d\Phi_{nw}(p'_{nw})$, and the expected capital gain of the unemployed native worker from finding a job through referrals is $\lambda_{nc} \int_{p'_{0nc}}^{\infty} [W_n - U_n] d\Phi_{nc}(p'_{nc})$. So the net expected capital gain of the unemployed native worker from finding a job is the sum of these two expressions. Note that after matching with the firm through formal channels only those unemployed native workers are hired whose productivity signal is higher than the lower bound of the productivity signal p'_{0nw}. While after matching with the firm through referrals only those unemployed native workers are hired whose productivity signal is higher than the lower bound of the productivity signal p'_{0nc}.

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

$$rU_i = b + \lambda_{fi} \int_{p'_{0iw}}^{\infty} [W_i - U_i] d\Phi_{iw}(p'_{iw}) + \lambda_{ic} \int_{p'_{0ic}}^{\infty} [W_i - U_i] d\Phi_{ic}(p'_{ic}) \tag{2.20}$$

The unemployed migrant worker also gets unemployment benefit b, and expects to move into employment through formal channels at job finding rate λ_{fi}. So the expected capital gain of the unemployed migrant worker from finding a job through formal channels is $\lambda_{fi} \int_{p'_{0iw}}^{\infty} [W_i - U_i] d\Phi_{iw}(p'_{iw})$. While the unemployed migrant worker expects to move into employment through referrals at rate λ_{ic}. Then the expected capital gain of the unemployed native worker from finding a job through referrals is $\lambda_{ic} \int_{p'_{0ic}}^{\infty} [W_i - U_i] d\Phi_{ic}(p'_{ic})$.

Equation (2.19) and (2.20) can be simplified and rewritten as follows:

$$rU_n = b + (\lambda_{fn} P(p'_{nw} > p'_{0nw}) + \lambda_{nc} P(p'_{nc} > p'_{0nc})) \frac{w - rU_n}{r + \delta} \tag{2.21}$$

$$rU_i = b + (\lambda_{fi} P(p'_{iw} > p'_{0iw}) + \lambda_{ic} P(p'_{ic} > p'_{0ic})) \frac{w - rU_i}{r + \delta} \tag{2.22}$$

Where we assume that unemployed workers do not observe their own productivity.

Thus U_n and U_i do not depend on the individual productivity of the unemployed worker. So the $W_n - U_n$ and $W_i - U_i$ can be taken out of the integrals and substituted by the right-hand sides (henceforth RHS) of the second parts of the equations (2.17) and (2.18) respectively.

The unemployment rates of native and migrant workers are given by the differences between the flows into and out of the unemployment. $f - u_n$ employed native workers lose their job at rate δ, so the flow into the unemployment of the native workers is $\delta(f - u_n)$. While u_n unemployed native workers find a job through formal channels at job finding rate λ_{fn}. With probability $P(p'_{nw} > p'_{0nw})$ they have productivity signal higher than p'_{0nw}, and they move to the employment. So the flow out of the unemployment for the native workers through the formal channels is $\lambda_{fn}P(p'_{nw} > p'_{0nw})u_n$. The u_n unemployed native workers may also find a job through referrals at job finding rate λ_{nc}. With probability $P(p'_{nc} > p'_{0nc})$ they have higher productivity signal than p'_{0nc}, and they move to the employment. Thus, the total outflow from the unemployment for the native workers is $\lambda_{fn}P(p'_{nw} > p'_{0nw})u_n + \lambda_{nc}P(p'_{nc} > p'_{0nc})u_n$. Similar to the employed native workers, $1 - f - u_i$ employed migrant workers lose their job at rate δ, so the flow into the unemployment for the migrant workers is $\delta(1 - f - u_i)$. u_i unemployed migrant workers find a job through formal channels at job finding rate λ_{fi}. With probability $P(p'_{iw} > p'_{0iw})$ they have higher productivity signal than p'_{0iw}, and they move to the employment. So the flow out of the unemployment for the migrant workers through the formal channels is $\lambda_{fi}P(p'_{iw} > p'_{0iw})u_i$. The u_i unemployed migrant workers may also find a job through referrals at job finding rate λ_{ic}. With probability $P(p'_{ic} > p'_{0ic})$ they have higher productivity signal than p'_{0ic}, and they move to the employment. Thus, the total outflow from the unemployment for the migrant workers is $\lambda_{fi}P(p'_{iw} > p'_{0iw})u_i + \lambda_{ic}P(p'_{ic} > p'_{0ic})u_i$. Steady state equations for the unemployment rates of natives and migrants can be described by the following equations:

$$\dot{u_n} = \delta(f - u_n) - \lambda_{fn}P(p'_{nw} > p'_{0nw})u_n - \lambda_{nc}P(p'_{nc} > p'_{0nc})u_n = 0 \tag{2.23}$$

$$\dot{u_i} = \delta(1 - f - u_i) - \lambda_{fi}P(p'_{iw} > p'_{0iw})u_i - \lambda_{ic}P(p'_{ic} > p'_{0ic})u_i = 0 \tag{2.24}$$

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

$$\frac{\delta(f - u_n)}{u_n} = \lambda_{fn}P(p'_{nw} > p'_{0nw}) + \lambda_{nc}P(p'_{nc} > p'_{0nc}) \tag{2.25}$$

$$\frac{\delta(1 - f - u_i)}{u_i} = \lambda_{fi}P(p'_{iw} > p'_{0iw}) + \lambda_{ic}P(p'_{ic} > p'_{0ic}) \tag{2.26}$$

The expression on the RHS of the equation (2.25) appears in the equation (2.21) either. The same holds for the equations (2.26) and (2.22). If we substitute the corresponding expressions in the equations (2.21) and (2.22) with the LHS of the equations (2.25) and

(2.26) we can obtain the following expressions for the present discounted value of unemployed native and migrant workers:

$$U_n = \frac{b(r+\delta) + w\frac{\delta(f-u_n)}{u_n}}{r(r+\delta+\frac{\delta(f-u_n)}{u_n})} \tag{2.27}$$

$$U_i = \frac{b(r+\delta) + w(\frac{\delta(1-f-u_i)}{u_i})}{r(r+\delta+\frac{\delta(1-f-u_i)}{u_i})} \tag{2.28}$$

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

$$rE[J(p|p')] = E(p|p') - w - \delta(E[J(p|p')] - V) \tag{2.29}$$

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 (2.28) can be rewritten so that we get the net return of the job to the firm in the LHS of the equation:

$$E[J(p|p')] - V = \frac{E(p|p') - w - rV}{r+\delta} \tag{2.30}$$

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

$$\begin{aligned} rV = -c + \frac{m_n}{v}\int_{p'_{0nc}}^{\infty}[E_p(J(p|p'_{nc})) - V]d\Phi_{nc}(p'_{nc}) + q_n\int_{p'_{0nw}}^{\infty}[E_p(J(p|p'_{nw})) - V]d\Phi_{nw}(p'_{nw}) \\ + \frac{m_i}{v}\int_{p'_{0ic}}^{\infty}[E_p(J(p|p'_{ic})) - V]d\Phi_{ic}(p'_{ic}) + q_i\int_{p'_{0iw}}^{\infty}[E_p(J(p|p'_{iw})) - V]d\Phi_{iw}(p'_{iw}) = 0 \end{aligned} \tag{2.31}$$

Where the flow cost of the vacancy is c. The expected return of the vacancy to the firm consists of four parts, because the are two different groups of workers matched to the firm through two different channels. 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 bound for the groups are so that the $E(p|p'_0) - w = 0$, i.e. the RHS of the equation (2.30) 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 (2.6) to (2.9) for the expected productivity at the lower bounds and equalizing them to the wage:

$$E(p|p'_{0nw}) = \mu + \frac{\sigma^2}{\sigma^2+\sigma^2_\epsilon}(p'_{0nw}-\mu) = w, p'_{0nw} = (w-\mu)\frac{\sigma^2+\sigma^2_\epsilon}{\sigma^2} + \mu \tag{2.32}$$

$$E(p|p'_{0nc}) = \mu + s + \frac{\sigma^2}{\sigma^2+\sigma^2_\epsilon}(p'_{0nc}-\mu-s) = w, p'_{0nc} = (w-\mu-s)\frac{\sigma^2+\sigma^2_\epsilon}{\sigma^2} + \mu + s \tag{2.33}$$

$$E(p|p'_{0iw}) = \mu - d + \frac{\sigma^2}{\sigma^2+\sigma^2_\epsilon}(p'_{0iw}-\mu+d) = w, p'_{0iw} = (w-\mu+d)\frac{\sigma^2+\sigma^2_\epsilon}{\sigma^2} + \mu - d \tag{2.34}$$

$$E(p|p'_{0ic}) = \mu - d + s + \frac{\sigma^2}{\sigma^2+\sigma^2_\epsilon}(p'_{0ic}-\mu+d-s) = w, p'_{0ic} = (w-\mu+d-s)\frac{\sigma^2+\sigma^2_\epsilon}{\sigma^2} + \mu - d + s \tag{2.35}$$

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

$$fr_n = \frac{\lambda_{nc}P(p'_{nc} > p'_{0nc})}{\lambda_{fn}P(p'_{nw} > p'_{0nw}) + \lambda_{nc}P(p'_{nc} > p'_{0nc})} \tag{2.36}$$

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

$$fr_i = \frac{\lambda_{ic}P(p'_{ic} > p'_{0ic})}{\lambda_{fi}P(p'_{iw} > p'_{0iw}) + \lambda_{ic}P(p'_{ic} > p'_{0ic})} \tag{2.37}$$

2.3.4 Wage determination

As already mentioned above, the union and the firm bargain over the wage ex-ante. The union maximizes $((f-u_n)W_n + (1-f-u_i)W_i + u_nU_n + u_iU_i - \frac{b}{r})$, 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 two different groups of employed workers matched to the firms through two different channels. The expression is the sum of the expected returns from the two groups of employed workers matched to the firms through two different channels. Since fr_n fraction of the $f-u_n$ employed native workers are matched to the firm through referrals, the expected return from the native workers who found job through referrals is $(f-u_n)fr_n\int_{p'_{0nc}}^{\infty}[E_p(J(p|p'_{nc}))-V]d\Phi_{nc}(p'_{nc})$. $1-fr_n$ fraction of the $f-u_n$ employed native workers are matched to the firm through formal channels. The expected return of a firm from a native worker who found job through formal channels is $\int_{p'_{0nw}}^{\infty}[E_p(J(p|p'_{nw}))-V]d\Phi_{nw}(p'_{nw})$. Thus the expected return of a firm from native workers who found job through formal channels is $(f-u_n)(1-fr_n)\int_{p'_{0nw}}^{\infty}[E_p(J(p|p'_{nw}))-V]d\Phi_{nw}(p'_{nw})$. fr_i fraction of the $1-f-u_i$ employed migrant workers are matched to the firm through referrals, so the expected return from the migrant workers who found job through referrals is $(1-f-u_i)fr_i\int_{p'_{0ic}}^{\infty}[E_p(J(p|p'_{ic}))-V]d\Phi_{ic}(p'_{ic})$. Finally, $1-fr_i$ fraction of $1-f-u_i$ employed migrant workers found their job through formal channels, hence the expected return of a firm from migrant workers who found job through formal channels is $(1-f-u_i)(1-$

$fr_i) \int_{p'_{0iw}}^{\infty} [E_p(J(p|p'_{iw})) - V] d\Phi_{iw}(p'_{iw})$. The outside option for the firm is zero.

$$\left((f-u_n)W_n + (1-f-u_i)W_i + u_nU_n + u_iU_i - \frac{b}{r}\right)^B$$
$$\left((f-u_n)fr_n \int_{p'_{0nc}}^{\infty} [E_p(J(p|p'_{nc})) - V] d\Phi_{nc}(p'_{nc})\right.$$
$$+ (f-u_n)(1-fr_n) \int_{p'_{0nw}}^{\infty} [E_p(J(p|p'_{nw})) - V] d\Phi_{nw}(p'_{nw})$$
$$+ (1-f-u_i)fr_i \int_{p'_{0ic}}^{\infty} [E_p(J(p|p'_{ic})) - V] d\Phi_{ic}(p'_{ic})$$
$$\left. + (1-f-u_i)(1-fr_i) \int_{p'_{0iw}}^{\infty} [E_p(J(p|p'_{iw})) - V] d\Phi_{iw}(p'_{iw})\right)^{1-B} \longrightarrow \max_w \quad (2.38)$$

Where B is the bargaining power of the union. This approach is an extension of the canonical model of "Right-to-Manage" (see Cahuc et al. (2014), Part Two, Chapter 7, section 3.2.) The model described above is used to numerically calculate some values of variables in the model and find an explanation of the 7.26% difference in the probabilities of using referrals when finding a job between the natives and migrants.

2.4 Numerical example

First, we discuss the choice of the values of the exogenous variables described in the Table 2.5. The mean of the workers' true productivity μ is normalized to 1. Stops (2016) and Dengler et al. (2016) use as a dependent variable the natural logarithm of the number of matches in the German labour market. The total factor productivity of the matching function is calculated using the estimation results of Stops (2016), and the average of the calculated values is around 0.75. The average total factor productivity of the matching function is around 1.05 when using the estimation results of Dengler et al. (2016). The total factor productivity of the matching function λ_0 is chosen 0.9, which is the average of the calculated values. Petrongolo and Pissarides (2001a) 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. Unemployment benefit b is equal to 0.7, which is close to the average in the literature. Shimer (2005a) sets the value of unemployment benefit to 0.4. Stupnytska (2015) uses the value of 0.5 for b. Hall and Milgrom (2008) get a larger value of 0.71, while Hagedorn and Manovskii (2008) set b equal to 0.955 in their benchmark calibration. The values of the flow cost of the vacancy, interest rate and job destruction rate are the values used by Stupnytska (2015).

Variable	Value	Explanation. Source.
μ	1	Mean of the workers' true productivity. Normalization.
λ_0	0.9	Total factor productivity of the matching function. Stops (2016), Dengler et al. (2016).
η	0.5	Elasticity with respect to the unemployment in the matching function Petrongolo and Pissarides (2001a).
f	0.8783	Fraction of the native workers. SOEP data from 2002 to 2008.
h_n	90	Number of the contacts of native workers. Stupnytska (2015).
h_i	50	Number of the contacts of migrant workers. Stupnytska (2015).
b	0.7	Unemployment benefit. Average in the literature.
σ^2	0.1	Variance of the true productivity. Own calculations.
σ_ϵ^2	0.1	Variance of the error term. Own calculations.
c	0.5	Flow cost of the vacancy. Stupnytska (2015).
γ_n	0.8783	Level of homophily between the social contacts of a native worker SOEP data from 2002 to 2008.
γ_i	0.7	Level of homophily between the social contacts of a migrant worker Titzmann and Silbereisen (2009).
r	0.01	Interest rate. Stupnytska (2015).
δ	0.2	Job destruction rate. Stupnytska (2015).

Table 2.5: 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 native worker we use the fraction of natives assuming natives form contacts randomly.

The model is calibrated using the (2.23) and (2.24) steady state equations for the unemployment rates of natives and migrants, Free Entry condition described in the equation (2.31), and the fractions of migrant and native workers who found their job through referrals from the equations (2.37) and (2.36). 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 fractions of natives and migrants who found job through referrals and the unemployment rates of natives and migrants are

Variable	Value	Explanation
s	0.1852	Average prod. difference between the workers with and without a referral
d	0.2471	Average productivity difference between the native and migrant workers
a	0.0616	Vacancy arrival rate
v	1.4785	Vacancy rate
B	0.839	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	0.2753	Fraction of natives who found job through referrals
fr_i	0.3479	Fraction of migrants who found job through referrals

Table 2.6: Calibration results

plugged into the five above mentioned equations. The system of the equations is numerically solved for s, d, a, v and B. The calibration results are presented in the Table 2.6. Further, Table 2.7 shows some of the values of the endogenous variables.

Variable	Value	Explanation
λ_{fn}	4.3033	Job finding rate in the formal market for natives
λ_{fi}	8.4922	Job finding rate in the formal market for migrants
λ_{nc}	0.9389	Job finding rate through referrals for natives
λ_{ic}	1.3784	Job finding rate through referrals for migrants
w	1.0430	Wage
q_n	0.1882	Job filling rate in the formal market for natives
q_i	0.0954	Job filling rate in the formal market for migrants
m_n	0.0607	Number of matches through referrals for natives
m_i	0.0229	Number of matches through referrals for migrants

Table 2.7: Values of the endogenous variables

Figure 2.1 shows the value of the objective function of the union and the firm at different values of wage w during the bargaining process. The green line is the objective function of the union. We can see that the objective function of the union reaches its maximum when wage equals to 1.084. The blue line is the objective function of the firm, which is decreasing in wage w.

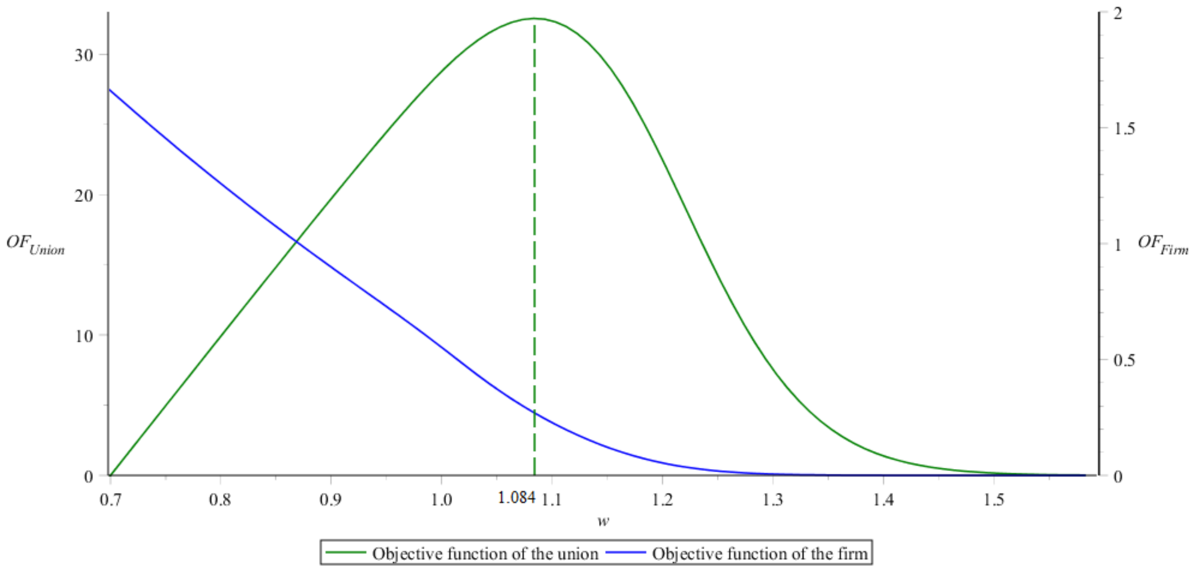

Figure 2.1: Objective functions of the union and the firm.

Figure 2.2 shows the value of the objective function at different values of wage w during the bargaining process. We can see that the value of the objective function reaches its maximum when the wage equals to 1.043. As expected, compared to the objective function of the union, the overall objective function is maximized at a lower wage.

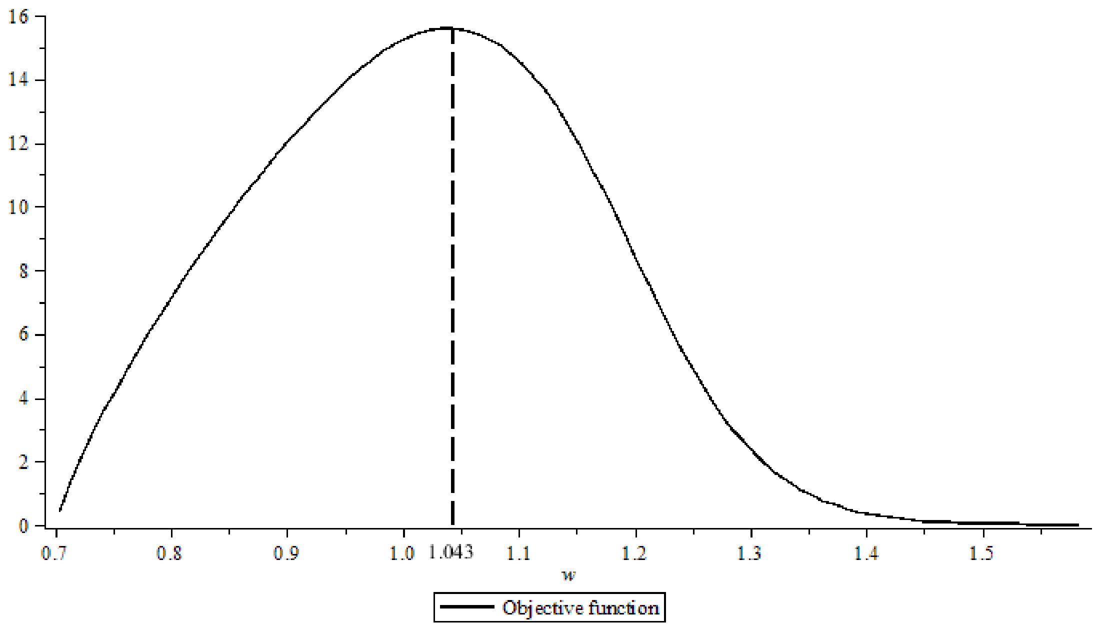

Figure 2.2: Objective function of the Nash bargaining process.

Figure 2.3 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

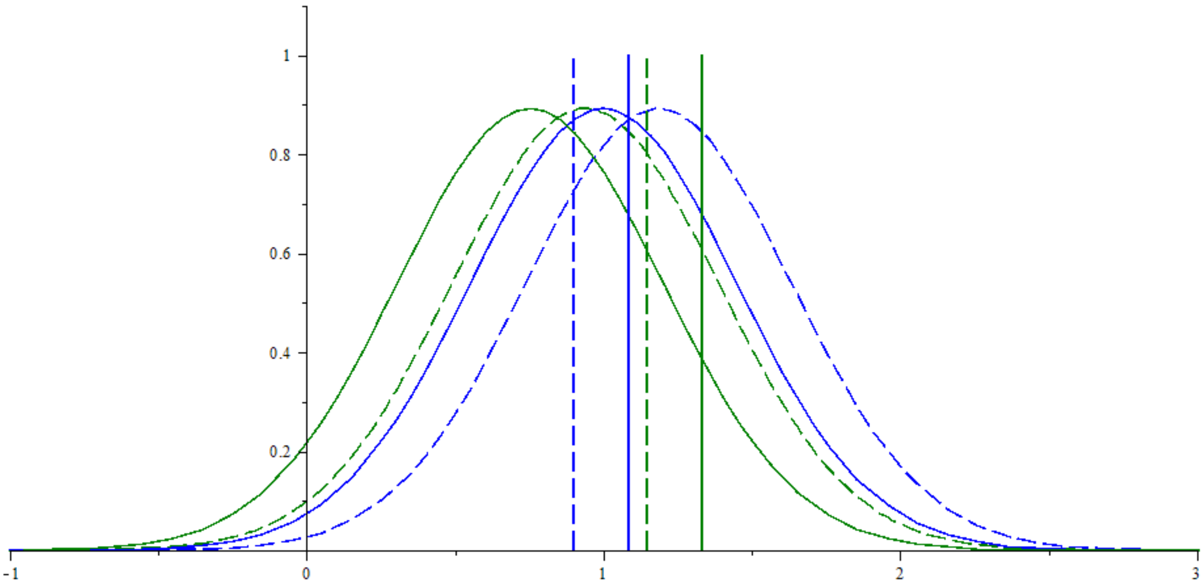

Figure 2.3: Probability density functions and the lower bounds of the productivity signals.

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 probabilities and the values of the lower bounds of the four groups are reported in the Table 2.8.

p'_{0nw}	$P(p'_{nw} > p'_{0nw})$	p'_{0nc}	$P(p'_{nc} > p'_{0nc})$	p'_{0iw}	$P(p'_{iw} > p'_{0iw})$	p'_{0ic}	$P(p'_{ic} > p'_{0ic})$
1.086	0.424	0.901	0.738	1.333	0.097	1.148	0.319

Table 2.8: Values of lower bounds and probabilities to be hired after the match

Table 2.8 shows that the probability that the native workers are hired after they are matched with the firm through referrals is 0.7377, which is approximately 1.74 times more than the probability for the native workers matched through formal channels. The probability for the native workers matched through formal channels is 0.4237. The probability that the migrant workers are hired after they are matched with the firm through referrals is 0.3195, 3.29 times more than the probability for the migrant workers matched through formal channels, which equals to 0.0972. 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 7.26% difference in the probabilities of using referrals when finding a job between the natives and migrants.

2.5 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 7.26% 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 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 0.2471. Moreover, average productivity difference between the workers with and without referrals is also positive and equals to 0.1852. 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 2.8 show that the probability that the native workers are hired after they are matched with the firm through referrals is approximately 1.74 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 3.29 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.

Chapter 3

Immigration, Social Networks and Occupational Mismatch

3.1 Introduction

In this study we investigate the link between the methods of job search that workers use to find employment and the probability of occupational mismatch in the new job. According to multiple empirical studies the most common search methods include private and public employment agencies, direct applications to job advertisements posted in newspapers and internet as well as help from friends and relatives. Following the literature we define referral hiring via the network of friends and relatives as an informal search channel, whereas employment agencies and direct applications form a formal channel of job search. The primary question that we address in this study is whether both search channels are equally efficient in generating good matches. By good matches we mean jobs in the original occupation corresponding to the professional training and education of the worker. Empirical evidence shows that changing the occupation is often associated with lower wages and higher job instability (Wolbers (2003), Allen and De Weert (2007), Robst (2007)), thus new jobs involving occupational mismatch can be seen as low quality matches. Moreover, we analyze if the efficiency of the search channel is the same for different demographic groups, with a particular focus on differences between native and immigrant workers.

In our empirical estimation we use data from the German Socio-Economic Panel (SOEP) over the period 2000-2014. This is a household survey which includes detailed information about worker characteristics, the job search method which was used to find the job as well as some characteristics of the employer. The data also includes subjective evaluation of the worker if the current job corresponds to his/her professional training or not. We use this information to form a proxy variable for occupational mismatch. In the first step, we document that referral hiring via social networks is the most frequent single channel of generating jobs in Germany. But there are large differences in the utilization of this channel between native and foreign workers. Whereas 31.5% of German workers found their current job by recommendation, this fraction is 43.8% for immigrant workers living in Germany.

Note, however, that this difference doesn't fully compensate immigrant workers for the lower chances of finding jobs via the formal channel, so the average risk of unemployment is higher for immigrants. This finding is particularly important in the view of the result by Bentolila et al. (2010) that referral hiring via social networks often generates mismatch between occupational choices of workers and their professional training. Intuitively, this means that social networks often serve as a method of last resort for workers and allows them to avoid unemployment at the cost of lower wages in the mismatch occupation. Hence we ask a question whether a more intensive utilization of social networks can lead to more frequent occupational mismatch of immigrant workers?

To address this question we develop a theoretical search and matching model with two ethnic groups of workers (natives N and immigrants I), two search channels (formal and referral hiring) and two occupations. This is a second step in our research. Half of the workers have initial professional training in occupation A but they can also perform jobs in occupation B, which is associated with occupational mismatch. The situation is symmetric in the two occupations. Depending on the ethnic background (N or I) and professional training (A or B) there are four distinct worker groups in the model. Thus workers in a given group have social links within their own group but also with workers in the other three groups. When modeling social networks we take into account ethnic and professional homophily. Intuitively, this means that foreign (native) workers have a larger fraction of other foreign (native) workers in their social network. Following the definition by Jackson (2010) ethnic bias in the formation of social networks can be characterized as homophily by choice since workers with similar ethnic background have common language, traditions and history. In contrast, occupational bias in the formation of social networks is homophily by opportunity since workers from the same profession/occupation are likely to have studied or worked together in the past.

In our model firms with open positions either make their vacancies public and try to fill the job in a formal way or contact one of the employees in their occupation and ask this employee to recommend a friend. In this latter case the position can be filled by referral hiring as workers transmit vacancy information to their unemployed social contacts. Whereas referral hiring is modeled endogenously, the processes of formal hiring and job destruction are based on the exogenous transition rates. In the numerical example of the model we choose these transition rates by targeting some of the key endogenous variables in the model, such as the unemployment rates and the rates of referral hiring observed in the German data. In order to incorporate the evidence by Bentolila et al. (2010) we normalize the rate of occupational mismatch generated by the formal channel to zero and investigate relative differences in the mismatch rates of native and immigrant workers generated by social networks. Our model predicts that higher rates of referral hiring among immigrants produce more frequent occupational mismatch of the immigrant population. One condition for this result is that the gap in the job destruction rates between native and immigrant workers is not too large which is satisfied for a realistic parameter setting motivated by the data. From a theoretical perspective the gap in mismatch rates strongly depends on the degree of professional homophily characterizing social networks and on the incidence of referrals but is not sensitive to the overall network size.

In the third step we validate the result by Bentolila et al. (2010) with the German dataset (SOEP) and test the main prediction of our model. Our data reveals that referral hiring is associated with the highest rate of occupational mismatch among all channels in Germany. It is equal to 53.5%, whereas the rate of occupational mismatch associated with direct applications to a vacancies advertised in internet is equal to 31.4%. Even though these rates are based on subjective evaluations of workers there is a remarkable difference in the observed frequencies which confirms the result by Bentolila et al. (2010) and the underlying setup of our theoretical model. Further, the data shows that immigrant workers have a significantly higher probability of occupational mismatch (57%) than native workers (42%) which is compatible with the main prediction of our model. However, it is not only this negative link between being a foreigner and the probability of a good match that we want to test, but the underlying mechanism of the model based on the search channel. So we included both binary variables for the immigration status and for referral hiring as a successful search channel into the logistic panel regression with a probability of a good match as predicted outcome. Our estimation shows that the negative marginal effect of the immigration indicator is reduced once we control for the job search channel which confirms our predictions that at least a part of the higher probability of mismatch in the group of foreign workers is explained by more frequent referral hiring.

In the last step we quantify the contribution of more intensive network hiring in the group of foreign workers to higher rates of occupational mismatch in this group. In order to achieve this goal we perform a Blinder-Oaxaca decomposition of differences in the occupational mismatch between native and foreign workers based on the linear probability model. Differences in the endowments between natives and foreigners including the job search channel jointly explain about a half of the gap in the mismatch rates between the two groups, that is 7.6% out of 15.5%. Most of this endowment effect (6.7% out of 7.6%) is explained by the lower education of foreign workers and by the industry effects. Intuitively, this means that foreign workers are overrepresented in industries with lower education and associated with higher rates of occupational mismatch such as transportation and trade. Nevertheless, the remaining 0.9% of the endowment effect is due to the less efficient search channels used by foreign workers. Thus the fact that foreign workers rely intensively on the support from their social networks contributes significantly to the higher rate of occupational mismatch of foreigners even though this effect is quantitatively smaller than the effect of classical explanatory factors such as education and industry.

3.1.1 Related literature

This paper is closely related to the literature on referral hiring, occupational mismatch and immigration. Even though bilateral relationships between these three components are reasonably well investigated, our study is a first theoretical and empirical attempt analyzing an integral relationship between all three components.

First, we contribute to the literature on referral hiring and match quality. Here a positive effect of referrals on match quality is highlighted by Montgomery (1991), Kugler (2003), Dustmann et al. (2016) and Galenianos (2013). The seminal study by Montgomery (1991) finds that 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, the workers hired through referrals are paid higher wages. The result is driven by the fact that social contacts tend to occur among workers with similar characteristics (homophily by ability), and that a worker will refer only well-qualified applicants, since his/her reputation is at stake. Whereas, Dustmann et al. (2016) distinguish between informal and formal search methods and build a model of ethnic networks. They predict that the probability of a minority worker from a particular ethnic group to be hired is positively related to the share of existing minority workers from that group in the firm. According to them workers hired through informal search methods initially get higher wages since the match-specific productivity is more uncertain when using formal methods, rather than informal methods. Kugler (2003) argues that employers which use informal methods in hiring are enabled to reduce their monitoring cost, and to pay lower efficiency wages because referees exert peer pressure on the referred workers. As a result, well-connected workers are matched to well-paid jobs.

Although most of the studies find that referrals increase the probability for the worker to be hired, Pistaferri (1999), Addison and Portugal (2002), Bentolila et al. (2010) and Zaharieva (2018) find negative wage effect of referrals. Our results are inline with the findings highlighted by Bentolila et al. (2010) for the United States. Even though social contacts reduce unemployment duration by about 1-3 months, they are associated with wage discounts of at least 2.5% due to occupational mismatch. This evidence reveals a trade-off from using social contacts in the job search: even though social contacts lead faster to new jobs and allow workers to leave unemployment, these jobs are more likely to be associated with occupational mismatch and lower wages. Pellizzari (2010) uses data from the European Community Household Panel (ECHP) and finds that in the European Union premiums and penalties to finding jobs through personal contacts are equally frequent and are of about the same size. Furthermore, he argues that wage penalties may be a result of mismatching, since they disappear with tenure. The advantage of our data compared to Bentolila et al. (2010) and Pellizzari (2010) is that it includes a direct indicator for occupational mismatch reported by the survey respondents. Furthermore, the goal of our study is to understand differences between native and immigrant workers in the use of social contacts and labour market outcomes, which was not done in the previous literature.

The studies by Zaharieva (2018) and Horvath (2014) develop theoretical models to study labour market outcomes of using social networks. Both studies introduce professional homophily into social networks which means that workers in a given profession have many friends and acquaintances from the same profession. Both authors document occupational mismatch being associated with the use of social networks in the job search. Moreover, the mismatch is decreasing with an increasing level of professional homophily. This is intuitive since a larger number of social contacts from the same profession make it more likely that a job referral will lead to a good match in this profession. Another two studies by Lancee (2016) and Alaverdyan (2018) incorporate ethnic homophily of social networks in their analysis which means that workers tend to have more friends of the same ethnic origin. To the best of our knowledge the model developed in the present paper is the first one that includes both dimensions of network homophily taking into account ethnic and professional characteristics of workers.

Second, our study is closely related to the literature on referral hiring and immigration. Immigrants are more likely to find their jobs through referrals compared to natives according to Drever and Hoffmeister (2008), Lancee (2016), Alaverdyan (2018). Other studies consider subgroups of immigrants from different countries of origin. For example, Ooka and Wellman (2006) investigate the importance of social networks in relation to the job search strategies of five immigrant groups living in Toronto. They find that Jewish immigrants have the highest rate of using personal contacts when searching for jobs (54%) followed by Italians (51%), Germans (45%), British (44%) and Ukrainians (40%). Elliot (2001) considers recent Latino immigrants to the United States. He finds that 81.1% of recent immigrants from this group were hired through the informal channel. The fraction is somewhat smaller for established immigrants (more than 5 years since arrival to the US) and equal to 72.8%. It falls down to 61.9% for Latino individuals born in the US. For comparison, the fraction of native US nationals finding jobs via the informal channel is 51.1%. These results indicate that referral hiring is a particularly important job search channel for recent immigrants in the United States but its importance declines with time as immigrant workers learn the local language and assimilate in the destination country.

Battu et al. (2011) find a similar assimilation effect of immigrant workers in the United Kingdom. They provide evidence that the less assimilated the ethnic unemployed workers are the more likely they are to use their network as their main method of job search. Moreover, they report that ethnic workers who obtained their current job as a result of their personal network are in a lower level job as a result. Again this indicates the fact that faster accession to jobs provided by social networks comes along with a wage penalty and worse job quality emphasized above. We complement this research direction by documenting that also in Germany the highest incidence of referrals is observed in the group of direct (first generation) immigrants (41.9%), followed by the indirect (second generation) immigrants (35.6%) and German nationals (30.3%). Moreover, we link these differences to the match quality of obtained jobs.

Third, we contribute to the debate on immigration and occupational mismatch. There is a vast literature on occupational mismatch distinguishing between vertical and horizontal mismatch. Vertical mismatch is observed when the worker is over- or underqualified for the occupation employed. While horizontal mismatch applies to the situation when the field of education of the worker does not correspond to the education required for the job (see Wolbers (2003), Allen and De Weert (2007) and Robst (2007)). Wolbers (2003) considers data on school graduates in Western European economies and finds that school-leavers from humanities, arts and agriculture are more likely to be mismatched than those from engineering, manufacturing, business and law. Robst (2007) finds similar results for college graduates in the United States and shows that 27-47% of workers in arts, social sciences, psychology, languages and biology are mismatched. He also reports that horizontal mismatch is associated with a wage loss of 10%. Also Iftikhar and Zaharieva (2019) use SOEP data and report substantial wage losses from mismatch in Germany focusing on the group of high skill workers. These findings suggest that occupational mismatch is an important labour market outcome with negative consequences on average even though it is possible that occasionally some workers gain higher salaries by changing the occupation.

More recent studies in this field compare the outcomes of native and immigrant workers. For example, Chiswick and Miller (2008) and Chiswick and Miller (2010) report lower returns to schooling for foreign-born workers compared to natives in the U.S. and Australia respectively and explain this outcome with low international transferability of immigrant's human capital skills implying more frequent skill mismatch of foreign-born workers. Aleksynska and Tritah (2013) consider a large set of European countries and find that immigrants are more likely to be both under- and overeducated than the native born for the jobs that they perform. However, immigrants' outcomes converge to those of the native born with the years of labor market experience. In our data we also observe this type of integration in the German labour market. Piracha and Vadean (2013) present an overview of this literature and show that the percentage of correctly matched immigrant employees is, for example, about 5.0% lower compared to native employees in Denmark and reaches up to 15.6% in the United States. The only exceptions are Finland and Italy, where the mismatch incidence seems to be higher for natives. They also point out that different measurement methods often lead to significantly different estimates of incidence rates. In particular, mismatch is more frequent when self-reported rather than when objective measures are used. Our empirical estimates for Germany are similar to the U.S. with the percentage of correctly matched immigrant employees 15.5% lower compared to natives. We contribute to this literature by explicitly comparing job search channels of workers and mismatch outcomes associated with these channels which was not done before. Moreover, we show that referral hiring generates occupational mismatch more frequently than other search strategies and it is this channel which is more often used by immigrant workers contributing to stronger occupational mismatch of this group.

Finally, there are several additional results that we obtain from the data. In particular, we document that educated workers are substantially less likely to use social contacts as intermediaries in the job search. Male workers are referred more often by their social contacts than female workers. This finding is generally consistent with the idea that women lack professional networks compared to men. It is also supported by the previous empirical research for the United States summarized in Marsden and Gorman (2001) and by Behtoui (2008) for women in Sweden. In addition, jobs in smaller companies are more frequently filled via social networks. This result is inline with the recent evidence in Rebien et al.

(2017) using German firm-level data.

The study proceeds as follows: in section 2 we describe the data and estimate regressions for the probability of finding a job via referrals. We use this empirical evidence to motivate our theoretical model which is developed and described in section 3. In section 4 we use empirical data to test new theoretical predictions of the model. More specifically, in this section we carry out the Blinder-Oaxaca decomposition of differences in the occupational mismatch rates between native and foreign workers. Section 5 concludes the paper.

3.2 Empirical evidence

In this section we describe our empirical data and analyze which factors can explain the risk of unemployment. We also explore the search channels used by workers to find employment. We use this empirical evidence to build up a job search model with two ethnic worker groups, two professional occupations and two different search channels: direct formal applications and referral hiring via social networks. The model is developed and presented in section 3.3. We also use predicted values of the key variables from this section to provide a realistic numerical example allowing us to illustrate the underlying economic mechanism of the model. In particular, we use the estimated unemployment rates and the fractions of workers who found their job through referrals by citizenship and migration background.

3.2.1 Estimation of unemployment rates

In this subsection we estimate unemployment rates for different worker groups by using empirical data from the German Socio-Economic Panel (SOEP). SOEP is a longitudinal study of households and individuals, which covers nearly 11,000 households, and about 30,000 individuals annually. Our sample covers data on 213592 individuals from SOEP 2000-2014. Among a wide range of questions regarding personal characteristics and employment data respondents are asked about their employment status and labour force status. The dependent variable $EMP_{i,t}$ is binary, and takes values $\{0, 1\}$ based on the answers to the above-mentioned questions. $EMP_{i,t}$ equals 1 if individual i is in full-time employment, marginal, regular or irregular part-time employment at time t. While $EMP_{i,t}$ equals 0 if individual i is non-working and registered unemployed at time t. Disabled individuals in sheltered employment, the individuals in military/community service, on maternity leave and in training program are excluded from the data. In addition, we exclude those non-working individuals which are older than 65, which are working past 7 days, those which have regular second job or occasional second job.

$MIG_{i,t}$ is a variable indicating the nationality of individuals. We define an individual to be foreign citizen if the person has foreign citizenship, and German citizen if the person has German citizenship. So, variable $MIG_{i,t}$ equals 1 if the i^{th} individual is a foreign citizen at time t, and it is equal to 0 if the i^{th} individual is a German citizen at time t. Additionally, $MIGBACK_{i,t}$ indicates the migration background of individuals based on their place of birth. If the respondent is born in another country, then the respondent is considered to have a direct migration background. If the respondent is born in Germany, but one of the respondent's parents has a migration background, then the respondent is considered to have an indirect migration background. While when there is no information about the respondent's migration background, then the respondent is classified as a German national.

Table 3.1: Percentage of unemployed individuals by citizenship\migration background.

Citizenship\ Migration background	Unemployed(%)	Unemployed	Employed	Total	Total(%)
Foreign Citizens	14.81%	2569	14772	17341	8.12%
German Citizens	7.86%	15421	180830	196251	91.88%
Direct migrants	13.30%	3784	24677	28461	13.32%
Indirect migrants	10.04%	1221	10941	12162	5.69%
German nationals	7.51%	12985	159984	172969	80.98%

According to the descriptive statistics presented in Table 3.1, 14.81% of foreign citizens are unemployed, compared to 7.86% for German citizens. While, 13.30% of direct migrants, 10.04% of indirect migrants, and 7.51% of German nationals are unemployed. So, the difference in unemployment rates between direct migrants and German nationals is higher than the difference between indirect migrants and German nationals. This might possibly be explained by partial assimilation of indirect migrants and better language skills, compared to direct migrants.

The descriptive statistics presented in Table 3.1 show that foreign citizens are more likely to be unemployed, but the reason may be due to different characteristics of the groups. To control for differences in the observable characteristics we regress $EMP_{i,t}$ on different variables sequentially adding the following variables to the regression equation. $EDU_{i,t}$ shows the amount of the i^{th} individual's education or training in years at time t computed by the SOEP (for detailed description see Helberger (1988) and Schwarze et al. (1991)). The values of $EDU_{i,t}$ range from 7 to 18. The i^{th} individual's age at time t is denoted by $AGE_{i,t}$. The dummy variable $FEMALE_{i,t}$ takes value 1 if the i^{th} individual is female at time t. The categorical variable $MARST_{i,t}$ shows the marital status of the i^{th} individual at time t. It has 5 categories: married/living with a partner, single, widowed, divorced, and separated (legally married). Another categorical variable $STATE_{i,t}$ indicates the German federal state in which the household of the i^{th} individual was located at the time of the survey. And finally, $NCHILD_{i,t}$ shows the number of persons in the household of the i^{th} individual under the age of 18 at time t. When the dependent variable is binary this study uses logistic regression model for estimations, and likelihood-ratio test to choose between regression equations. After adding each variable to the regression equation a likelihood-ratio test is conducted to see if the variable added contributes statistically significantly to the regression. The main estimation results of the regression equations are presented in Table 3.2. The detailed estimation results with the coefficients of all variables are presented in Table 3.13 in Appendix I.

Table 3.2: Employment rates: logistic regression

Variables	Dependent variable: EMP							
(1)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EDU	0.020***	0.019***	0.019***	0.019***	0.019***	0.019***	0.019***	0.018***
	(94.84)	(94.76)	(94.93)	(94.98)	(101.42)	(101.70)	(100.00)	(90.46)
AGE		0.00014**	0.00014**	-0.00064***	-0.00049***	-0.00050***	-0.00088***	-0.00095***
		(3.24)	(3.25)	(-13.47)	(-11.28)	(-11.55)	(-18.46)	(-19.98)
FEMALE			-0.0086***	-0.0063***	-0.0058***	-0.0059***	-0.0063***	-0.0069***
			(-8.70)	(-6.58)	(-6.86)	(-6.97)	(-7.36)	(-8.08)
MARST(Reference: Married)								
[2] Single				-0.058***	-0.045***	-0.045***	-0.058***	-0.061***
				(-33.51)	(-28.97)	(-29.03)	(-31.92)	(-33.02)
[3] Widowed				-0.017***	-0.014***	-0.014***	-0.016***	-0.017***
				(-4.90)	(-4.54)	(-4.51)	(-5.13)	(-5.42)
[4] Divorced				-0.050***	-0.043***	-0.043***	-0.045***	-0.047***
				(-25.75)	(-24.47)	(-24.52)	(-25.78)	(-26.46)
[5] Separated				-0.048***	-0.043***	-0.043***	-0.045***	-0.045***
				(-12.87)	(-12.53)	(-12.52)	(-13.09)	(-13.20)
NCHILD							-0.009***	-0.009***
							(-20.08)	(-19.07)
MIG								-0.043***
								(-19.57)
STATE					v	v	v	v
Survey year t						v	v	v
LR test(Prob> χ^2)		0.0012	0.00	0.00	0.00	0.00	0.00	0.00
Observations	213592	213592	213592	213592	213592	213592	213592	213592
Pseudo R^2	0.062	0.062	0.063	0.081	0.116	0.117	0.120	0.125

Table 3.2 reveals that education is positively associated with the employment probability. Also married workers are more likely to be employed. In contrast, being a female reduces the probability of employment. The negative and statistically significant coefficient of variable $MIG_{i,t}$ indicates that foreign citizens are less likely to be employed. The predicted probabilities of being employed for two otherwise-average individuals' are 94.84% for German citizens, and 90.55% for foreign citizens. So the risk of unemployment is 5.16% for the first group and 9.45% for the second group. We use these predicted values of the unemployment rates in the numerical example of the model in section 3.3. The results of the likelihood-ratio tests suggest that all the above-mentioned variables should be added to the regression equation. When variable $MIGBACK_{i,t}$ is added to the regression equation instead of $MIG_{i,t}$, the qualitative result doesn't change[1]. The predicted probabilities of being employed for otherwise-average individuals' from the three groups are the following: 95.27% for German nationals, 92.37% for indirect migrants and 90.53% for direct migrants. The predicted probability of being employed for indirect migrants is closer to the probability for German nationals, compared to direct migrants, which indicates some degree

[1]The coefficients for this regression are available on demand from the authors.

of assimilation. Note that in all regressions the predicted probabilities are estimated at the average values of control variables. Next we analyze the incidence of different search channels used by workers to find employment with a particular focus on referral hiring.

3.2.2 Estimation of referral hiring

The respondents of the SOEP survey who started their current job within the previous two years answer the question how they found their current job. One of the possible answers is that information about the job was provided by friends or relatives of the respondent. We classify these cases as referral hiring (informal channel). Other search channels such as the federal employment office, an advertisement in the internet or newspaper, a job-center (ARGE) and a private recruitment agency are classified as formal channels. The value of the corresponding dummy variable $REF_{i,t}$ equals 1 if the i^{th} individual found the job via a referral from some friend or relative, and it equals 0 if the i^{th} individual used a formal channel to find the job.

Table 3.3: Percentage of individuals who found their job through referrals by citizenship\migration background.

Citizenship\ Migration background	Found job through referrals(%)	Found job through		Total	Total(%)
		Referrals	Formal chan.		
Foreign Citizens	43.84%	648	830	1478	7.72%
German Citizens	31.48%	5562	12108	17670	92.28%
Direct migrants	41.91%	873	1210	2083	10.88%
Indirect migrants	35.58%	528	956	1484	7.75%
German nationals	30.86%	4809	10772	15581	81.37%

According to the descriptive statistics presented in Table 3.3, 43.84% of foreign citizens found their job through referrals, compared to 31.48% for German citizens. Following a different definition 41.91% of direct migrants, 35.58% of indirect migrants, and 30.86% of German nationals obtained help from their friends and relatives. So, the difference in the proportion of individuals who found their job through referrals between indirect migrants and German nationals is lower than the difference between direct migrants and German nationals.

In the next step $REF_{i,t}$ is regressed on a set of control variables to test if the differences in referral hiring are due to the different characteristics of the two groups. In addition to variables indicating the individuals' education, age, gender, state of residence, and survey year the following variables are sequentially added to the regression equation. $FSIZE_{i,t}$ is a categorical variable with four categories showing the size of the firm in which the i^{th} individual is employed at time t. The categories are: less than 20 employees, 20 to 200, 200 to 2000, and more than 2000 employees. Another categorical variable $IND_{i,t}$ indicates the industry of i^{th} individual at time t. $IND_{i,t}$ has 9 categories: Agriculture, Energy, Mining, Manufacturing, Construction, Trade, Transport, Bank/Insurance, and Services. The categorical variable $TOJCH_{i,t}$ has 5 categories and indicates which kind of job change preceded the current employment of individual i. The categories of $TOJCH_{i,t}$ are the following: first job, job after break, job with new employer, company taken over, changed job at the same firm. Last, the Standard International Socio-Economic Index of Occupational Status developed by Ganzeboom et al. (1992) is used to control for the occupational status. ISEI index reflects individual's socio-economic status based on information about this individual's income, education, and occupation. $ISEI_{i,t}$ index takes values in the range between 16 and 90.

To see if the independent variable contributes significantly to the regression a likelihood-ratio test was conducted for all new control variables. The main estimation results are presented in Table 3.4. While the detailed estimation results with the coefficients of all variables are presented in Table 3.14 in Appendix II.

Table 3.4: Estimation results of referral hiring.

Variables	Dependent variable: REF									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
EDU	-0.023***	-0.023***	-0.021***	-0.017***	-0.015***	-0.014***	-0.014***	-0.014***	-0.006***	-0.006***
	(-17.59)	(-17.44)	(-16.13)	(-12.62)	(-10.83)	(-9.69)	(-9.64)	(-9.76)	(-3.63)	(-3.62)
AGE		-0.00063*	-0.00062*	-0.00094**	-0.00090**	-0.00057	-0.00056	-0.00049	-0.00062	-0.00063
		(-2.11)	(-2.08)	(-3.13)	(-2.97)	(-1.74)	(-1.72)	(-1.48)	(-1.88)	(-1.89)
MIG			0.085***	0.085***	0.082***	0.077***	0.075***	0.077***	0.074***	0.072***
			(6.41)	(6.33)	(6.13)	(5.80)	(5.54)	(5.75)	(5.53)	(5.44)
FSIZE(Reference: GE 2000)										
[1] LT 20				0.154***	0.152***	0.110***	0.110***	0.108***	0.097***	0.098***
				(15.99)	(15.54)	(10.81)	(10.79)	(10.60)	(9.33)	(9.39)
[2] GE 20 LT 200				0.077***	0.076***	0.039***	0.040***	0.037***	0.030**	0.030**
				(8.03)	(7.82)	(3.89)	(3.94)	(3.69)	(2.93)	(2.91)
[3] GE 200 LT 2000				0.028**	0.031**	0.008	0.008	0.007	0.004	0.004
				(2.67)	(2.89)	(0.71)	(0.73)	(0.61)	(0.36)	(0.33)
IND					v	v	v	v	v	v
TOJCH(Reference: First job)										
Job After Break						-0.060***	-0.060***	-0.068***	-0.074***	-0.072***
						(-4.58)	(-4.57)	(-5.18)	(-5.60)	(-5.49)
Job With New Employer						0.030*	0.030*	0.034**	0.031*	0.031*
						(2.42)	(2.40)	(2.77)	(2.52)	(2.51)
Company Taken Over						-0.243***	-0.242***	-0.243***	-0.246***	-0.246***
						(-14.25)	(-14.23)	(-14.26)	(-14.49)	(-14.54)
Changed Job, Same Firm						-0.255***	-0.255***	-0.256***	-0.256***	-0.256***
						(-18.41)	(-18.41)	(-18.48)	(-18.36)	(-18.36)
STATE							v			
Survey year t								v	v	v
ISEI									-0.0022***	-0.0022***
									(-7.95)	(-8.00)
FEMALE										-0.0157*
										(-2.10)
LR test(Prob> χ^2)		0.0344	0.00	0.0275	0.00	0.00	0.5708	0.00	0.00	0.00
Observations	19148	19148	19148	19148	19148	19148	19148	19148	19148	19148
Pseudo R^2	0.013	0.014	0.015	0.028	0.030	0.058	0.058	0.060	0.062	0.062

t statistics in parentheses. Marginal effects for factor levels is the discrete change from the base level.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.4 shows that referral hiring is more important for less educated workers and it is more widespread in smaller firms. First employment and jobs with new employers are more likely to be generated by means of referral hiring. Moreover, the negative coefficient of the dummy variable $FEMALE_{i,t}$ indicates that female workers are less likely to be hired through referrals than male workers. The results of likelihood-ratio tests suggest that except $STATE_{i,t}$ all the above-mentioned variables should be added to the regression equation.

The positive and statistically significant coefficient of variable $MIG_{i,t}$ indicates that foreign citizens are more likely to find their jobs through referrals. The predicted probabilities of finding a job through referral for two otherwise-average individuals' are 29.72% for German citizens, and 36.96% for foreign citizens. We use these values in the numerical example of the model in section 3.3. When variable $MIGBACK_{i,t}$ is added to the regression equation instead of $MIG_{i,t}$ predicted probabilities of finding a job through referrals for otherwise-average individuals' from the three groups are the following: 29.26% for German nationals, 36.47% for direct migrants, and 32.36% for indirect migrants[2]. Thus, the predicted probability of finding a job through referrals for indirect migrants is closer to the probability for German nationals, compared to direct migrants.

In the next step we use this empirical evidence to develop a theoretical search and matching model capturing differences in the unemployment rates and job search strategies of native and foreign workers. We use this model to address a question if differences in the search strategies may contribute to differences in the match qualities between the two groups.

3.3 The model

In this section we develop a search and matching model with two occupations, two search channels (formal search and network referrals) and two ethnic worker groups (natives and foreigners). The model incorporates the fact that foreign workers rely more often on their social networks when searching for jobs which was documented in the previous section. It also allows for different unemployment rates of the two ethnic worker groups. The objective of developing this model is to analyze the impact of referral hiring on occupational mismatch of native and foreign workers. In addition, we use the model to understand the implications of other factors such as network characteristics and labour market properties for the link between network hiring and occupational mismatch.

Consider a model with two professional groups of infinitely lived risk neutral workers and two occupations. Workers of type A obtained training in occupation A, which is their primary occupation, but they can also work in occupation B, which is a mismatch occupation for them. In a similar way, occupation B is a primary occupation for type

[2]The coefficients for this regression are available on demand from the authors.

B workers, whereas there is mismatch if type B workers are employed in occupation A. Each group of workers is a continuum of measure 1. In each professional group there is a fraction h of foreign workers F and a fraction $1-h$ of native workers N. Hence there are four demographic groups in the economy $\{N, A\}$, $\{F, A\}$, $\{N, B\}$ and $\{F, B\}$.

Consider native type i individuals, $i = A, B$. Each person can be unemployed (u_N^i), employed and well matched in the original occupation (m_N^i) or mismatched and employed in another occupation (x_N^i). The same holds for foreign type i individuals with corresponding notation u_F^i, m_F^i and x_F^i, so we get:

$$u_N^i + m_N^i + x_N^i = 1 - h \qquad u_F^i + m_F^i + x_F^i = h$$

In addition, let e_j^i, $i = A, B$ and $j = N, F$ denote all employed workers of type j and profession i, both matched and mismatched, that is:

$$e_N^i = m_N^i + x_N^i \qquad e_F^i = m_F^i + x_F^i$$

Let v^A and v^B denote exogenous stocks of open vacancies in occupations A and B respectively. There are two channels of job search: formal applications and referrals via the social network (informal channel). Only unemployed workers are searching for a job, so there is no on-the-job search. We follow the assumption of Bentolila et al. (2010) and assume that workers always send their formal applications to vacancies in their original occupation. This assumption is based on the empirical evidence that social networks generate occupational mismatch more frequently than formal search. We verify this assumption for Germany in section 3.4. Even though in reality formal applications can also lead to mismatch, we normalize it to zero to investigate the relative difference in mismatch rates generated by the two search channels.

To simplify the model occupations A and B are assumed to be symmetric. Let λ_N and λ_F denote the job-finding rates of native and foreign workers via the formal channel in each of the two occupations. Variables δ_N and δ_F denote the job destruction rates of native and foreign workers in each of the two occupations. These rates do not depend on the way the worker found the job and do not depend on the occupation. Nevertheless, we allow for possible differences in the job stability of native and foreign workers. Since the focus of our study is on referral hiring we assume that the rates λ_N, λ_F, δ_N and δ_F are exogenously given. To model referral hiring let n denote the number of social contacts in the networks of workers. We assume that the network size n is the same for all individuals. Furthermore, social networks exhibit professional and ethnic homophily. A more detailed composition of social networks is described in the next subsection.

3.3.1 Social networks

Consider a native type A individual. This person has some social contacts within his/her group, let their number be denoted by n_{NN}^{AA}. In addition, this person knows some foreign workers from the same occupation, let their number be denoted by n_{NF}^{AA}. In the same way there are some links between this person and individuals in occupation B, let them be denoted by n_{NN}^{AB} and n_{NF}^{AB}. Here the former number stands for the links to native type B workers and the latter number for the links to foreign type B workers. So in general every native person of type A has contacts within each of the four demographic groups. Given that the total number of contacts for one person is denoted by n we get:

$$n_{NN}^{AA} + n_{NF}^{AA} + n_{NN}^{AB} + n_{NF}^{AB} = n$$

The composition of social networks is illustrated on figure 3.1. Next consider foreign type A workers. Their contacts within the group are denoted by n_{FF}^{AA} and their contacts with native type A workers are denoted by n_{FN}^{AA}. Variables n_{FN}^{AB} and n_{FF}^{AB} stand for the links to native and foreign workers in occupation B respectively, so we get:

$$n_{FN}^{AA} + n_{FF}^{AA} + n_{FN}^{AB} + n_{FF}^{AB} = n$$

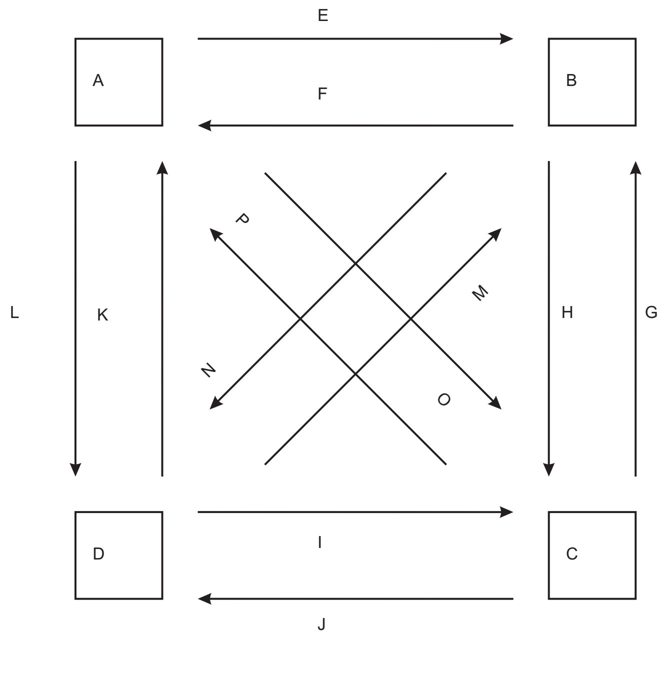

Figure 3.1: Composition of social networks

Social networks exhibit professional and ethnic homophily. In general, homophily refers to the fact that people are more prone to maintain relationships with others who are similar to themselves. There can be homophily by age, race, gender, religion, ethnicity or professional occupation and it is generally a robust observation in social networks (see McPherson et al. (2001) for an overview of research on homophily). The focus of this paper is on the latter two types of homophily by ethnicity and occupation. Jackson (2010) distinguishes between homophily due to opportunity and due to choice. In this respect, homophily by occupation is likely to arise due to the fact that workers with the same profession studied or worked together in the beginning of their career. Thus it is rather a limited opportunity of meeting workers from different professions which generates homophily rather than an explicit choice. In contrast, homophily by ethnicity is likely to be a choice outcome since workers with similar ethnicity/origin share common background, values and traditions which makes their communication easier.

Let $\gamma \in [0.5..1]$ denote the degree of professional homophily, identical for all workers. This means that every worker has a fraction γ of contacts in the same occupation and a fraction $1-\gamma$ of contacts in the other occupation. This means:

$$n_{NN}^{AA} + n_{NF}^{AA} = \gamma n \qquad n_{FN}^{AA} + n_{FF}^{AA} = \gamma n$$

In the extreme case when $\gamma = 1$ workers in different occupations are completely disconnected. The opposite case $\gamma = 0.5$ corresponds to random matching without homophily. This is due to the fact that both professional groups A and B are equally large.

In addition, social networks are characterized by ethnic homophily, let $\tau \geq h$ denote the fraction of foreign individuals in the network of a foreign person. So we get:

$$n_{FN}^{AA} = (1-\tau)\gamma n \qquad n_{FF}^{AA} = \tau\gamma n \qquad n_{FN}^{AB} = (1-\tau)(1-\gamma)n \qquad n_{FF}^{AB} = \tau(1-\gamma)n$$

This is the network composition of foreign type A workers parametrized by γ and τ. Furthermore, social networks should be balanced. The total number of links from native individuals of type A to foreigners of type A given by $(1-h)n_{NF}^{AA}$ should be the same as the total number of links from foreign individuals of type A to natives of type A given by hn_{FN}^{AA}. Moreover, the total number of links from native individuals of type B to foreign individuals of type A, that is $(1-h)n_{NF}^{BA}$, should be the same as the number of links from foreign individuals of type A to native individuals of type B given by hn_{FN}^{AB}. This means:

$$(1-h)n_{NF}^{AA} = hn_{FN}^{AA} \qquad (1-h)n_{NF}^{BA} = hn_{FN}^{AB}$$

Inserting $n_{FN}^{AA} = (1-\tau)\gamma n$ and $n_{FN}^{AB} = (1-\tau)(1-\gamma)n$ we get:

$$\begin{aligned} n_{NF}^{AA} &= \tfrac{h(1-\tau)\gamma n}{1-h} & n_{NN}^{AA} &= \frac{(1-2h+h\tau)\gamma n}{1-h} \\ n_{NF}^{BA} &= \tfrac{h(1-\tau)(1-\gamma)n}{1-h} & n_{NN}^{BA} &= \frac{(1-2h+h\tau)(1-\gamma)n}{1-h} \end{aligned}$$

This is a consistent network composition of native type A workers parametrized by γ and τ. To obtain the last equation we used the fact that the two occupations are symmetric and $n_{NN}^{BA} + n_{NF}^{BA} = (1-\gamma)n$. These equations show that if $\tau \geq h$, that is the fraction of foreign contacts in the networks of foreigners τ is larger than their population fraction h, then it also holds that the fraction of native contacts in the networks of natives $(1-2h+h\tau)/(1-h)$ is larger than their population fraction $1-h$ because $(1-2h+h\tau)/(1-h) > 1-h$. Thus ethnic homophily should be seen as a two-sided process.

Note an important special case when $\tau = h$. This is a situation when foreign and native workers are randomly mixed and create links with each other. So there is no ethnic homophily and both groups have a fraction h of foreigners in their networks ($n_{NF}^{AA} = n_{FF}^{AA} = h\gamma n$) and a fraction $1 - h$ of natives ($n_{NN}^{AA} = n_{FN}^{AA} = (1-h)\gamma n$).

Further, symmetry between the two occupations implies the same composition of social networks for type B workers, so that $n_{FN}^{BB} = n_{FN}^{AA}$, $n_{FF}^{BB} = n_{FF}^{AA}$, $n_{FN}^{BA} = n_{FN}^{AB}$, $n_{FF}^{BA} = n_{FF}^{AB}$ and $n_{NN}^{BB} = n_{NN}^{AA}$, $n_{NF}^{BB} = n_{NF}^{AA}$, $n_{NN}^{BA} = n_{NN}^{AB}$, $n_{NF}^{BA} = n_{NF}^{AB}$. In order to illustrate the composition of social networks in our model we complement this subsection with a small example.

Example of network composition: Let $\gamma = \tau = 0.6$, $n = 50$ and $h = 0.2$. This means that the fraction of foreign workers in the economy is 20%. Then we get the following composition of networks:

$$\begin{aligned} n_{FN}^{AA} &= 12 \qquad n_{FF}^{AA} = 18 \qquad n_{FN}^{AB} = 8 \qquad n_{FF}^{AB} = 12 \\ n_{NF}^{AA} &= 3 \qquad n_{NN}^{AA} = 27 \qquad n_{NF}^{AB} = 2 \qquad n_{NN}^{AB} = 18 \end{aligned}$$

Both foreign and native workers know 30 contacts in their own occupation and 20 contacts in the other occupation. This is because $\gamma = 30/50 = 0.6$. But the ethnic composition of social networks is very different. Whereas the networks of native workers are very extreme with only 3 links to foreign workers and 27 links to other native workers in their occupation, the networks of foreign workers are more equal with 12 links to native workers and 18 links to other foreign workers in the same occupation. The reason for this effect is twofold. On the one hand, foreign workers are a minority in the labour market which implies that native workers are much less likely to meet a foreigner and create a contact than the other way round. Even if matching was balanced with respect to ethnic belonging we would expect that native workers know only $0.2 \cdot 30 = 6$ foreign workers and 24 other natives in their occupation. On the other hand, the distribution becomes even more extreme with ethnic homophily, since $\tau = 0.6$. This example shows that even though we use identical parameters γ, τ and n for native and foreign workers, the actual networks generated by these parameters are very different between these two groups.

As we emphasized in the introduction, there are many empirical studies showing that referrals from social contacts are important in the job search process. Our example reveals that the situation of native and foreign workers is asymmetric in this respect. Whereas foreign workers are likely to receive important vacancy information from their native and foreign friends, foreign contacts are unlikely to be an important source of job-related information for native workers. In the next subsection we analyze more specifically how vacancy information is transmitted in the market and derive referral probabilities for all demographic groups.

3.3.2 Transition rates

In this subsection we derive endogenous network transition rates from unemployment to jobs for all worker groups. Recall that λ_N and λ_F are the exogenous job-finding rates via the formal channel. By assumption formal applications always lead to jobs in the original occupation. In contrast, network referrals can lead to both types of jobs in the original occupation and in the mismatch occupation. Let μ_N^{AA} and μ_F^{AA} denote the network job-finding rates of native and foreign workers of type A in occupation A respectively. In addition, let μ_N^{AB} and μ_F^{AB} denote network job-finding rates leading to mismatch jobs in occupation B. The structure of worker flows and the corresponding job-finding rates are presented on figure 3.2. The network job-finding rates are illustrated by the dashed arrows.

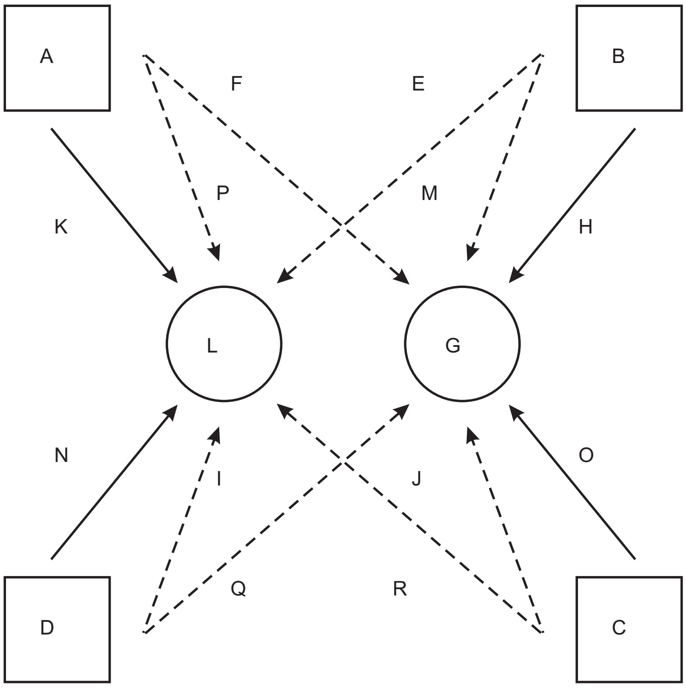

Figure 3.2: Structure of the labour market

Consider vacancies in occupation A. With an exogenous probability s firms with open vacancies in this occupation contact one of the incumbent type A employees and ask this employee to recommend a friend for the open position. It is intuitive to think that firms only ask those employees who are properly matched to the job, these are workers m_N^A and m_F^A. So with probability $m_j^A/(m_N^A+m_F^A)$ the firm contacts the employee with ethnic origin $j = N, F$.

Further we assume that every contacted type A employee is first considering his/her unemployed friends of the same type. Only if all type A friends are employed the person considers unemployed contacts of type B. Some rationale for this assumption could be that well matched type A workers in occupation A are more productive than mismatched type B workers. Among type A contacts the person has n_{jN}^{AA} native friends and n_{jF}^{AA} foreign friends. So with probability $[e_N^A/(1-h)]^{n_{jN}^{AA}}$ all native friends of this employee are employed and with probability $[e_F^A/h]^{n_{jF}^{AA}}$ all foreign friends of this employee are also employed. This means that $1 - [e_N^A/(1-h)]^{n_{jN}^{AA}}[e_F^{AA}/h]^{n_{jF}^{AA}}$ is a probability that this employee can recommend at least one unemployed friend searching for the job. So the number of network matches between type A vacancies and type A native workers recommended by the employee $j = N, F$ is:

$$M_{jN}^{AA} = sv^A \frac{m_j^A}{m_N^A + m_F^A}\left(1 - \left[\frac{e_N^A}{1-h}\right]^{n_{jN}^{AA}}\left[\frac{e_F^A}{h}\right]^{n_{jF}^{AA}}\right)\frac{n_{jN}^{AA}\cdot\frac{u_N^A}{1-h}}{n_{jN}^{AA}\cdot\frac{u_N^A}{1-h} + n_{jF}^{AA}\cdot\frac{u_F^A}{h}}$$

where the last term is a probability that a randomly chosen unemployed type A friend of the employee is native. In the special case without ethnic homophily ($\tau = h$) we get $n_{jN}^{AA} = (1-h)\gamma n$ and $n_{jF}^{AA} = h\gamma n$, $j = N, F$. So the above expression can be simplified as:

$$M_{jN}^{AA} = sv^A \frac{m_j^A}{m_N^A + m_F^A}\left(1 - \left[\frac{e_N^A}{1-h}\right]^{(1-h)\gamma n}\left[\frac{e_F^A}{h}\right]^{h\gamma n}\right)\frac{u_N^A}{u_N^A + u_F^A}$$

In a similar way, the number of network matches between type A vacancies and type A foreign workers recommended by the employee $j = N, F$ is given by:

$$M_{jF}^{AA} = sv_A \frac{m_j^A}{m_N^A + m_F^A}\left(1 - \left[\frac{e_N^A}{1-h}\right]^{n_{jN}^{AA}}\left[\frac{e_F^A}{h}\right]^{n_{jF}^{AA}}\right)\frac{n_{jF}^{AA}\cdot\frac{u_F^A}{h}}{n_{jN}^{AA}\cdot\frac{u_N^A}{1-h} + n_{jF}^{AA}\cdot\frac{u_F^A}{h}}$$

where the last term is a probability that a randomly chosen unemployed type A friend of employee j is a foreigner. We can see that the total number of good matches between type A vacancies and type A unemployed native workers per unit time is given by $M_{NN}^{AA} + M_{FN}^{AA}$. In addition, the total number of good matches between type A vacancies and type A unemployed foreign workers per unit time is $M_{NF}^{AA} + M_{FF}^{AA}$. Given that the stocks of searching unemployed native and foreign workers are u_N^A and u_F^A the network transition rates into the original occupation for native and foreign workers can be calculated as:

$$\mu_N^{AA} = \frac{M_{NN}^{AA} + M_{FN}^{AA}}{u_N^A} \qquad \mu_F^{AA} = \frac{M_{NF}^{AA} + M_{FF}^{AA}}{u_F^A}$$

That is the flow probability of finding a job by recommendation in the primary occupation is given by the ratio between the total number of good matches in this occupation and the total number of searching workers separately for each ethnic group. Here we account for all possible situations including cases when native workers are recommended by their foreign friends and vice verse. Lemma 1 presents our results for the special case when $\tau = h$.

Lemma 1: *Network transition rates within the original occupation are the same for native and foreign workers in the absence of ethnic homophily ($\tau = h$), that is* $\mu^{AA} \equiv$

μ_N^{AA}=μ_F^{AA} *and:*

$$\mu^{AA} = \frac{sv_A}{u_N^A + u_F^A}\left(1 - \left[\frac{e_N^A}{1-h}\right]^{(1-h)\gamma n}\left[\frac{e_F^A}{h}\right]^{h\gamma n}\right)$$

The same is true in occupation B, that is $\mu_N^{BB} = \mu_F^{BB}$.
Proof: Appendix.

In the special case when social networks do not exhibit ethnic homophily and $\tau = h$ the composition of networks is the same among native and foreign workers. This means that both groups have a fraction h of foreigners and a fraction $1-h$ of natives among their occupation-specific contacts. So the probability of hearing about a job via the network in their primary occupation is the same for both groups.

Next consider occupation B. With the same probability s firms with open vacancies v_B ask one of the incumbent type B employees to recommend a friend. Recall that workers of type B have native (n_{jN}^{BB}) and foreign friends (n_{jF}^{BB}) in their occupation. This gives rise to matches M_{jN}^{BB} and M_{jF}^{BB} in a similar way as above. However, with probability $[e_N^B/(1-h)]^{n_{jN}^{BB}}[e_F^B/h]^{n_{jF}^{BB}}$ the employee doesn't have any unemployed type B friends. Recall that this employee also has native (n_{jN}^{BA}) and foreign friends (n_{jF}^{BA}) in occupation A. So the employee is considering unemployed type A friends. With probability $(1 - [e_N^A/(1-h)]^{n_{jN}^{BA}}[e_F^A/h]^{n_{jF}^{BA}})$ the employee knows at least one unemployed type A person who is searching for a job, so a new match is created. Let M_{jN}^{AB} denote the number of matches between type A native workers recommended by their type B friends with ethnic origin $j = F, N$:

$$M_{jN}^{AB} = \underbrace{\frac{sv_B \cdot m_j^B}{m_N^B + m_F^B}}_{(1)}\underbrace{\left[\frac{e_N^B}{1-h}\right]^{n_{jN}^{BB}}\left[\frac{e_F^B}{h}\right]^{n_{jF}^{BB}}}_{(2)}\underbrace{\left(1 - \left[\frac{e_N^A}{1-h}\right]^{n_{jN}^{BA}}\left[\frac{e_F^A}{h}\right]^{n_{jF}^{BA}}\right)}_{(3)}\underbrace{\frac{n_{jN}^{BA}\cdot\frac{u_N^A}{1-h}}{n_{jN}^{BA}\cdot\frac{u_N^A}{1-h} + n_{jF}^{BA}\cdot\frac{u_F^A}{h}}}_{(4)}$$

Here the first term is the probability that the firm is asking a type B employee with ethnic origin $j = N, F$ to recommend a friend. The second term corresponds to the probability that this employee doesn't have any unemployed type B friends. The third term is the probability that this employee knows at least one unemployed type A friend. And finally the last term is the probability that a randomly chosen unemployed type A friend of the employee is native.

In the special case without ethnic homophily ($\tau = h$) we know that $n_{jN}^{BA} = (1-h)(1-\gamma)n$ and $n_{jF}^{BA} = h(1-\gamma)n$. So the above expression can be written as:

$$M_{jN}^{AB} = \frac{sv_B \cdot m_j^B}{m_N^B + m_F^B}\left[\frac{e_N^B}{1-h}\right]^{(1-h)\gamma n}\left[\frac{e_F^B}{h}\right]^{h\gamma n}\left(1 - \left[\frac{e_N^A}{1-h}\right]^{(1-h)(1-\gamma)n}\left[\frac{e_F^A}{h}\right]^{h(1-\gamma)n}\right)\frac{u_N^A}{u_N^A + u_F^A}$$

Finally, the number of network matches between type B vacancies and type A foreign workers recommended by the employee $j = N, F$ is:

$$M_{jF}^{AB} = \frac{sv_B \cdot m_j^B}{m_N^B + m_F^B}\left[\frac{e_N^B}{1-h}\right]^{n_{jN}^{BB}}\left[\frac{e_F^B}{h}\right]^{n_{jF}^{BB}}\left(1 - \left[\frac{e_N^A}{1-h}\right]^{n_{jN}^{BA}}\left[\frac{e_F^A}{h}\right]^{n_{jF}^{BA}}\right)\frac{n_{jF}^{BA} \cdot \frac{u_F^A}{h}}{n_{jN}^{BA} \cdot \frac{u_N^A}{1-h} + n_{jF}^{BA} \cdot \frac{u_F^A}{h}}$$

where the last term is the probability that a randomly chosen unemployed type A friend of employee B is a foreigner. Given the number of matches, the network transition rates into the mismatch occupation for native and foreign workers are given by:

$$\mu_N^{AB} = \frac{M_{NN}^{AB} + M_{FN}^{AB}}{u_N^A} \qquad \mu_F^{AB} = \frac{M_{NF}^{AB} + M_{FF}^{AB}}{u_F^A}$$

Note here that both native and foreign social contacts can potentially lead to the mismatch job. Transition rates for type B workers μ_N^{BB}, μ_F^{BB}, μ_N^{BA} and μ_F^{BA} can be found symmetrically. Lemma 2 provides a summary of our results on the mismatch transition rates in the special case when $\tau = h$.

Lemma 2: *Network transition rates to the mismatch occupation are the same for native and foreign workers in the absence of ethnic homophily ($\tau = h$), that is $\mu^{AB} \equiv \mu_N^{AB} = \mu_F^{AB}$ and:*

$$\mu^{AB} = \frac{sv_B}{u_N^A + u_F^A}\left[\frac{e_N^B}{1-h}\right]^{(1-h)\gamma n}\left[\frac{e_F^B}{h}\right]^{h\gamma n}\left(1 - \left[\frac{e_N^A}{1-h}\right]^{(1-h)(1-\gamma)n}\left[\frac{e_F^A}{h}\right]^{h(1-\gamma)n}\right)$$

The same is true in occupation B, that is $\mu_N^{BA} = \mu_F^{BA}$.
Proof: similar to lemma 1.

Lemma 2 shows that if there are no differences in the composition of social networks between native and foreign workers and everyone has a population fraction h of foreign friends and $1-h$ of native friends in the network, then there are no differences in the mismatch transition rates between the two ethnic groups.

3.3.3 Equilibrium

In this subsection we analyze the dynamics of unemployment and matched employment for all worker groups and characterize the steady state of the model. The dynamics of unemployment u_N^A and matched employment m_N^A for native type A workers can be written as:

$$\begin{aligned}\dot{u}_N^A &= \delta_N(1 - h - u_N^A) - u_N^A(\lambda_N + \mu_N^{AA} + \mu_N^{AB}) \\ \dot{m}_N^A &= (\lambda_N + \mu_N^{AA})u_N^A - \delta_N m_N^A\end{aligned}$$

Here $\delta_N(1 - h - u_N^A)$ corresponds to employed type A workers losing jobs at rate δ_N, so it is the inflow into unemployment for native type A workers. At the same time the term $u_N^A(\lambda_N + \mu_N^{AA} + \mu_N^{AB})$ is the outflow of these workers from unemployment. It reflects the fact that there are three possibilities of finding a job: by means of a formal application at rate λ_N and with a help of friends/relatives at rate $\mu_N^{AA} + \mu_N^{AB}$. In the second equation the term $(\lambda_N + \mu_N^{AA})u_N^A$ corresponds to native type A workers finding jobs in their primary occupation, while $\delta_N m_N^A$ is the outflow of workers from this group due to job losses.

We have two similar equations for foreign workers:

$$\begin{aligned}
\dot{u}_F^A &= \delta_F(h - u_F^A) - u_F^A(\lambda_F + \mu_F^{AA} + \mu_F^{AB}) = 0 \\
\dot{m}_F^A &= (\lambda_F + \mu_F^{AA})u_F^A - \delta_F m_F^A = 0
\end{aligned}$$

In the steady state the outflow of workers from a given state should be equal to the inflow of workers into this state, so we set $\dot{u}_N^A = 0$, $\dot{m}_N^A = 0$, $\dot{u}_F^A = 0$ and $\dot{m}_F^A = 0$. So the steady-state distributions of workers across the three states are given by:

$$\begin{aligned}
u_F^A &= \frac{\delta_F h}{\delta_F + \lambda_F + \mu_F^{AA} + \mu_F^{AB}} \qquad m_F^A = \frac{(\lambda_F + \mu_F^{AA})h}{\delta_F + \lambda_F + \mu_F^{AA} + \mu_F^{AB}} \qquad x_F^A = h - u_F^A - m_F^A \\
u_N^A &= \frac{\delta_N(1-h)}{\delta_N + \lambda_N + \mu_N^{AA} + \mu_N^{AB}} \qquad m_N^A = \frac{(\lambda_N + \mu_N^{AA})(1-h)}{\delta_N + \lambda_N + \mu_N^{AA} + \mu_N^{AB}} \qquad x_N^A = 1 - h - u_N^A - m_N^A
\end{aligned} \tag{3.1}$$

Consider the simplified case without ethnic homophily, that is $\tau = h$. From lemmas 1 and 2 we know that the network transition rates in this case are the same for native and foreign workers, so that $\mu^{AA} = \mu_N^{AA} = \mu_F^{AA}$ and $\mu^{AB} = \mu_N^{AB} = \mu_F^{AB}$. From the empirical evidence presented in section 3.2 we also know that foreign workers rely more often on their social networks when searching for jobs, so the fraction of network hires is higher for foreign workers:

$$R_N = \frac{(\mu^{AA} + \mu^{AB})}{(\lambda_N + \mu^{AA} + \mu^{AB})} < \frac{(\mu^{AA} + \mu^{AB})}{(\lambda_F + \mu^{AA} + \mu^{AB})} = R_F$$

In our model we can capture this evidence by setting $\lambda_N > \lambda_F$. Intuitively, this means the following. If foreign workers face larger difficulties in the formal job search then referrals via social networks become a more important employment generating channel for foreign workers compared to natives. Several explanations for $\lambda_N > \lambda_F$ could be that there is more uncertainty associated with foreign training and education, worse language proficiency of foreigners and/or discrimination against ethnic minorities. Next we compare the mismatch rates of the two worker groups and see that:

$$\frac{x_N^A}{1-h} = \frac{\mu^{AB}}{\delta_N + \lambda_N + \mu^{AA} + \mu^{AB}} < \frac{x_F^A}{h} = \frac{\mu^{AB}}{\delta_F + \lambda_F + \mu^{AA} + \mu^{AB}} \quad \text{if} \quad \delta_N + \lambda_N > \delta_F + \lambda_F$$

This condition requires that $\delta_F - \delta_N < \lambda_N - \lambda_F$. Thus if the difference in the job destruction rates is not too large, then our *model predicts higher mismatch rates of foreign workers compared to natives.* There are two underlying processes that generate this prediction. On the one hand, empirical evidence from section 3.2 shows that network referrals are more important for foreign workers compared to natives. On the other hand, we incorporate the empirical evidence from Bentolila et al. (2010) that referral hiring leads more often to mismatch jobs compared to the formal search channel. Our model shows that a combination of these processes leads to the fact that foreign workers are more often mismatched in the equilibrium than native workers.

The above prediction is derived for the special case when $\tau = h$. In order to understand the situation in the more realistic case with ethnic homophily in the next subsection we set parameters to those observed in the German data and perform a detailed numerical analysis of model properties.

3.3.4 Numerical results

In this subsection we analyze model predictions in the more general case when social networks exhibit some degree of ethnic homophily. For this purpose we choose values of the exogenous parameters inline with existing empirical research. We also target several empirical variables reported in section 3.2. Given that the two sectors are symmetric we set $v = v^A = v^B$. Further note that the search intensity of firms s and the vacancy rate v are inseparable in the model and can only be determined as a product sv. From now on we consider sv as a single parameter. With this simplification the vector of exogenous parameters used in the model includes $\{\lambda_N, \lambda_F, \delta_N, \delta_F, sv, \tau, \gamma, n, h\}$.

Iftikhar and Zaharieva (2019) analyzed the size of foreign population in Germany over the period 2005-2016. They find that even though the fraction of foreign citizens was below 10% in Germany in this period, the fraction of individuals with immigration background was 18.2% in 2005 and it increased to 19.7% in 2013. Given that social networks are likely to evolve along the ethnic background rather than formal citizenship we set $h = 0.2$. Further this study shows that the average job duration of native workers in Germany was stable in the considered period and equal to 12 years. Given that the standard time unit in search and matching models is 1 quarter, we set $\delta_N = 0.02$, which corresponds to the average job duration of native workers equal to $1/0.02 = 50$ quarters. The average job duration for immigrant workers is substantially lower and close to 10 years. So we set $\delta_F = 0.03$ to capture the difference. Intuitively, this means that the jobs of foreign and immigrant workers are less stable compared to native workers.

We do not observe the size and homophily of social networks in labour market statistics. Cingano and Rosolia (2012) report that the median number of social connections between individuals in Italy is about 32. Glitz (2017) reports a comparable number for Germany with approximately 43 social contacts. In related theoretical studies Stupnytska and Zaharieva (2017) use 40 as the average network size, while it is 50 in Cahuc and Fontaine (2009). Zaharieva (2018) shows that the optimal diversification of social networks between two occupations strongly depends on the unemployment benefits and the mismatch wage relative to the wage in the primary occupation. Lower unemployment benefits and higher mismatch wages make social contacts outside the primary occupation more valuable and the optimal homophily parameter is low and close to 0.6 in this case. For this study we set $n = 30$ and $\gamma = 0.6$ as a starting point of the numerical investigation but we also perform comparative statics analysis with respect to both parameters and summarize the implications of the model for $\gamma \in [0.5..1]$ and $n \in [30..50]$.

In order to determine the remaining 4 parameters $\{\lambda_N, \lambda_F, sv, \tau\}$ we use our results from section 3.2 and target the following 4 endogenous variables: $u_N/(1-h) = 0.052$, $u_F/h = 0.094$, $R_N = 0.297$ and $R_F = 0.370$. Due to the symmetry assumption we use the same values in both occupations. These endogenous variables show that the unemployment rate of foreign/migrant workers is higher than the unemployment rate of native workers. Moreover, native workers rely less often on their social networks. Recall that R_j, $j = N, F$ is the fraction of referral hires out of new matches, which is given by:

$$R_N = \frac{(\mu_N^{AA} + \mu_N^{AB})u_N^A}{(\lambda_N + \mu_N^{AA} + \mu_N^{AB})u_N^A} \qquad R_F = \frac{(\mu_F^{AA} + \mu_F^{AB})u_F^A}{(\lambda_F + \mu_F^{AA} + \mu_F^{AB})u_F^A}$$

Using these two expressions and equations (3.1) for the equilibrium unemployment rates we find values of parameters $\{\lambda_N, \lambda_F, sv, \tau\}$ summarized in Table 3.5.

Table 3.5: Exogenous parameters and target variables

Parameter	Value	Target and Source
λ_N	0.256	Unemployment rate $u_N/(1-h) = 0.052$, SOEP
λ_F	0.182	Unemployment rate $u_F/h = 0.094$, SOEP
sv	0.008	Fraction of network hires $R_N = 0.297$, SOEP
τ	0.290	Fraction of network hires $R_F = 0.370$, SOEP

We can see that $\lambda_N = 0.256 > \lambda_F = 0.182$. This means that small differences in the job destruction rates between native and foreign workers ($\delta_N = 0.02 < \delta_F = 0.03$) are alone not sufficient to generate empirically observed differences in the unemployment rates between these two groups. So we can conclude that higher unemployment rates of foreign and immigrant workers in Germany are not only due to the lower stability of jobs occupied by the latter group but also due to lower chances of being hired upon a formal application. This result is inline with the experimental evidence presented in Kaas and Manger (2012). Moreover, we can see that $\tau = 0.290 > h = 0.2$. This means that social networks compatible with empirical evidence exhibit a moderate degree of ethnic homophily in Germany. Note that the average fraction of foreigners in the networks of native workers is $h(1-\tau)/(1-h) = 0.1775$, that is 17.75%. The equilibrium values of endogenous variables for our parameter choices are presented in Table 3.6.

Table 3.6 shows that the mismatch probability of natives $x_N^A/(1-h)$ is equal to 5.7% and it is lower compared to 6.8% for foreign workers. This numerical finding confirms our previous prediction that larger dependence of foreign workers on their social networks leads to more frequent mismatch of foreigners. We have already shown this in the special case when $\tau = h$ but it also holds in the more realistic case with ethnic homophily ($\tau > h$).

Table 3.6: Equilibrium values of endogenous variables

Native workers				Foreign workers			
Variable	Value	Variable	Value	Variable	Value	Variable	Value
$u_N^A/(1-h)$	0.052	μ_N^{AA}	0.086	u_F^A/h	0.094	μ_F^{AA}	0.085
$m_N^A/(1-h)$	0.891	μ_N^{AB}	0.022	m_F^A/h	0.838	μ_F^{AB}	0.022
$x_N^A/(1-h)$	0.057	R_N^A	0.297	x_F^A/h	0.068	R_F^A	0.370

In the next step we perform comparative statics analysis with respect to the compound parameter sv. Parameter s is driving the intensity of referral hiring in the model, if $s = 0$ firms don't use referrals to hire workers, in contrast, when s is large referral hiring dominates the formal search channel.

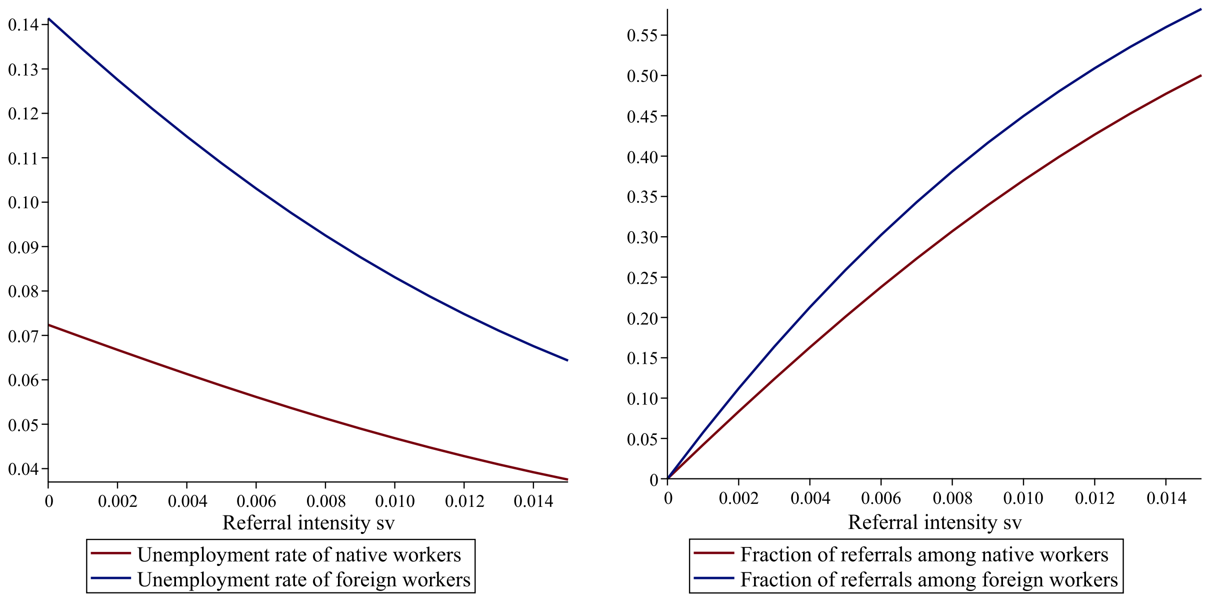

Figure 3.3: Left panel: Unemployment rates of native and foreign workers ($u_N/(1-h)$ and u_F/h) in the benchmark setting. Right panel: Fractions of network hires for native and foreign workers (R_N and R_F) in the benchmark setting

Our results are presented on figure 3.3. The left panel shows changes in the unemployment rates of the two ethnic groups. Finding jobs becomes easier for both groups when s is increasing. For example, both unemployment rates are two times smaller when $sv = 0.015$ compared to the case without referral hiring $sv = 0$. Even though the relative change is similar, the absolute drop in the unemployment rate of foreign workers is more pronounced compared to natives. The right panel of this figure shows changes in the fraction of referral hires R_N and R_F. Since formal applications of foreign workers are less successful compared to natives ($\lambda_F < \lambda_N$) informal hiring via networks becomes more important for foreigners. So we can see that $R_F > R_N$ for all realistic values of sv. To some extent referral hiring is a channel compensating the disadvantaged group for lower employment chances associated with formal applications.

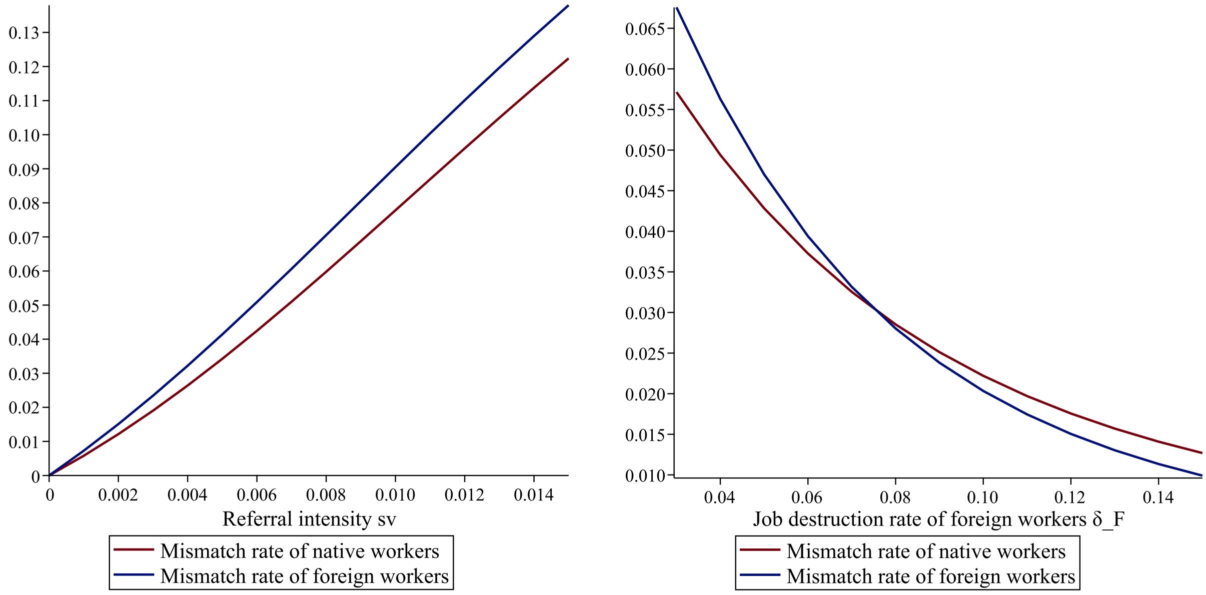

Figure 3.4: Left panel: Mismatch rates of native and foreign workers ($x_N^A/(1-h)$ and x_F^A/h), benchmark. Right panel: Mismatch rates of native and foreign workers ($x_N^A/(1-h)$ and x_F^A/h) for different values of δ_F

The left panel of figure 3.4 shows changes in the mismatch rates of the two ethnic groups. The fraction of mismatched foreign workers is higher than the fraction of mismatched native workers for all values of sv and the relative difference is increasing with more intensive referral hiring. Note that both rates start at zero, this is due to the normalization of mismatch to 0 in the absence of network hiring.

In section 3.3.3 we considered a simplified case without ethnic homophily and proved that foreign workers are more often mismatched if $\delta_F - \delta_N < \lambda_N - \lambda_F$. Note that this condition holds for the chosen parameter values. In order to understand the importance of this condition also in the more general case of ethnic homophily we increase parameter δ_F and illustrate the corresponding changes in both mismatch rates on the right panel of figure 3.4. We can see that with extreme values of δ_F the model may generate situations when the mismatch rate of native workers is higher than the mismatch of foreigners. If δ_F is extremely high than the jobs of foreign workers are very unstable and their unemployment rate is increasing very rapidly with the higher job destruction rate. In this situation very few foreign workers are employed in matched or mismatched employment as most of them are unemployed, so it may even happen that native workers are more often mismatched. However, this situation is not compatible with the realistic parameter values of δ_F.

Finally, we perform comparative statics analysis with respect to parameters γ and n since our empirical data is not sufficient to determine their values. Our results are illustrated on figure 3.5. We can see that the gap in the mismatch rates of foreign and native workers is decreasing with higher values of occupation homophily γ. This is intuitive since higher values of γ imply larger occupational segregation of workers, so the mismatch rates of both groups decrease and fall down to 0 when $\gamma = 1$. This is the case of complete occupational segregation. At the same time changes in the size of social networks n don't have strong implications for the relative difference in the mismatch rates of the two worker groups.

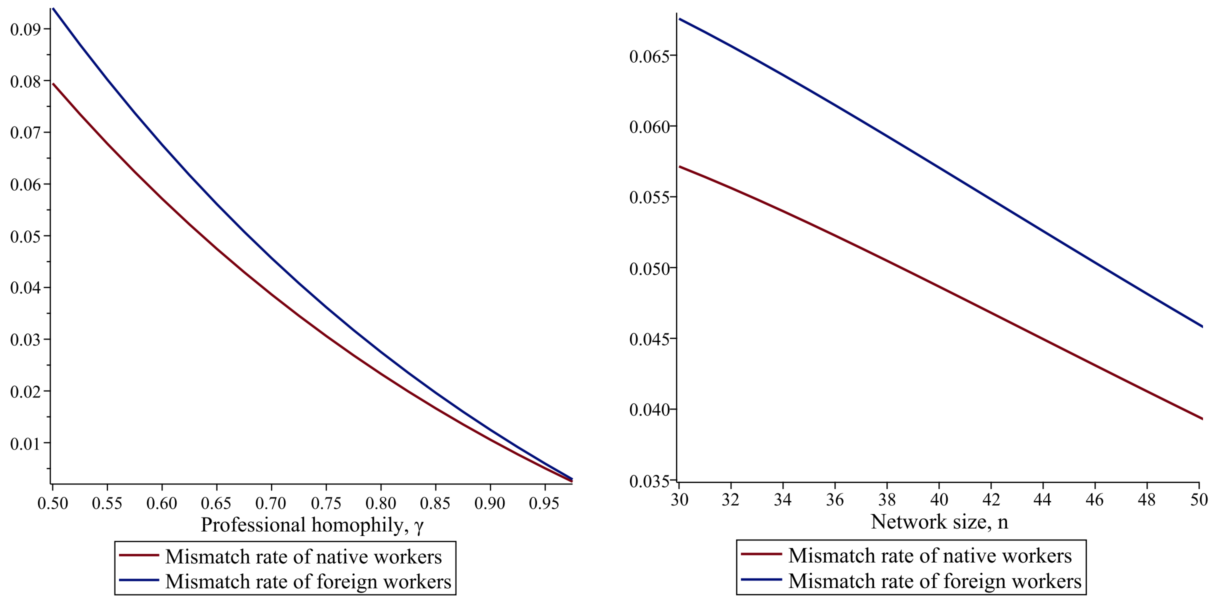

Figure 3.5: Left panel: Mismatch rates of native and foreign workers for different values of γ. Right panel: Mismatch rates of native and foreign workers for different values of n

To sum up, our theoretical analysis suggests that stronger reliance of foreign workers on referral hiring could be one of the reasons contributing to stronger occupational mismatch of foreigners compared to native workers. In the next section we continue our empirical analysis and test this theoretical prediction. We also test the underlying assumption of our model that referral hiring generates more occupational mismatch than formal search methods suggested by Bentolila et al. (2010).

3.4 Empirical testing

In this section we estimate the probabilities of occupational mismatch for different worker groups and discuss our findings. The main goal of our empirical analysis is to find answers to the following questions. Do the social networks generate more occupational mismatch compared to the formal search channels? Are foreign workers more likely to be mismatched compared to German workers? If yes, how much of the gap in mismatch rates between the two groups can be explained by stronger utilization of social networks by foreign workers?

3.4.1 Estimation of occupational mismatch

First, let us define occupational mismatch. The respondents who found their current job within the previous two years answer the question if they were educated of trained for their current position. The corresponding binary variable $MATCH_{i,t}$ takes value 1, if the i^{th} person answers that his or her position is the same as the profession for which he or she was educated or trained, thus the person is considered to be well matched. $MATCH_{i,t}$ takes value 0, if the i^{th} respondent is mismatched at time t. The respondents who are currently in training or have no previous training, are dropped from the sample. As a result descriptive statistics presented in Table 3.7 below is slightly different from the descriptive statistics presented in Table 3.3.

Descriptive statistics presented in Table 3.7 shows that foreign citizens are 15.51% more likely to be mismatched compared to German citizens. Furthermore, 56.33% of direct migrants are mismatched, while 42.00% of German nationals and 42.34% of indirect migrants are mismatched. The numbers for referral hiring have slightly changed due to the smaller sample size compared to section 3.2 but the qualitative conclusion is the same. So, migrants are more likely to find a job through referrals, and to be mismatched.

Next we investigate the job search channels in more details. The categorical variable $CHAN_{i,t}$ shows the channel through which individual i found his or her current job at time t. Workers are considered to have found their job through public employment agency if they respond that they found their current job through Employment Office, Job-Center, or Personal Service Agentur. They are considered to have found their job through other search channels if they respond that they found their current job by applying on chance, returned to former employer, or found a job through other search channels. The corresponding descriptive statistics are presented in Table 3.7. This table shows that referral hiring is a single most import search channel generating jobs in Germany, followed by newspapers, public employment agencies and direct applications in internet.

Table 3.7: Descriptive statistics of $MATCH_{i,t}$, $REF_{i,t}$, and $CHAN_{i,t}$ by citizenship\migration background.

	German citizens	Foreign citizens	German national	Direct migrants	Indirect migrants	Overall
MATCH						
Yes	57.56%	42.05%	58.00%	43.67%	57.66%	56.56%
No	42.44%	57.95%	42.00%	56.33%	42.34%	43.44%
REF						
Formal channels	69.57%	58.05%	70.07%	59.14%	67.54%	68.82%
Referrals	30.43%	41.95%	29.93%	40.86%	32.46%	31.18%
CHAN						
Public emp. agency	9.41%	10.73%	9.46%	10.68%	8.24%	9.50%
Private emp. agency	1.27%	1.66%	1.21%	1.92%	1.45%	1.30%
Newspaper	12.65%	14.15%	12.72%	13.11%	12.50%	12.74%
Internet	7.90%	4.59%	7.74%	7.03%	7.95%	7.68%
Referrals	30.43%	41.95%	29.93%	40.86%	32.46%	31.18%
Other	38.34%	26.93%	38.94%	26.41%	37.40%	37.59%
Observations	14754	1025	13183	1564	1032	15779
Percentage	93.50%	6.50%	83.55%	9.91%	6.54%	100%

Descriptive statistics for the control variables are presented in Table 3.15 in Appendix III. Besides statistics about the overall sample, Table 3.15 includes descriptive statistics separately for German citizens, foreign citizens, German nationals, direct and indirect migrants, to better understand the differences between these groups.

Table 3.8 shows that for all worker groups referrals lead most often to mismatch compared to all other search channels. Moreover, finding a job through the public employment agency leads to the second lowest percentage of good matches among the search channels. In contrast, finding a job through internet leads to the lowest percentage of mismatches for all the groups except indirect migrants. To sum up, our descriptive statistics shows that foreign citizens are more likely to be mismatched compared to German citizens, and compared to other search channels, referrals lead more often to occupational mismatch. Also, referrals reduce the probability of a good match for all groups, but relatively more so for foreign citizens.

Table 3.8: Descriptive statistics of $MATCH_{i,t}$ by search channels for different worker groups.

	Public emp. agency	Private emp. agency	Newspaper	Internet	Referrals	Other	Formal channels	Referrals	Overall
Yes	47.97%	55.12%	53.61%	68.56%	46.46%	65.69%	61.13%	46.46%	56.56%
No	52.03%	44.88%	46.39%	31.44%	53.54%	34.31%	38.87%	53.54%	43.44%
MATCH: German citizens									
Yes	48.52%	55.85%	54.07%	68.58%	47.73%	66.53%	**61.87%**	**47.73%**	57.56%
No	51.48%	44.15%	45.93%	31.42%	52.27%	33.47%	38.13%	52.27%	42.44%
MATCH: Foreign citizens									
Yes	40.91%	47.06%	47.59%	68.09%	33.26%	48.55%	**48.40%**	**33.26%**	42.05%
No	59.09%	52.94%	52.41%	31.91%	66.74%	34.31%	51.60%	66.74%	57.95%
MATCH: German nationals									
Yes	48.44%	53.75%	54.44%	69.61%	48.18%	66.86%	62.20%	48.18%	58.00%
No	51.56%	46.25%	45.56%	30.39%	51.82%	33.14%	37.80%	51.82%	42.00%
MATCH: Direct migrants									
Yes	40.72%	53.33%	44.88%	63.64%	35.68%	50.61%	49.19%	35.68%	43.67%
No	59.28%	46.67%	55.12%	36.36%	64.32%	49.39%	50.81%	64.32%	56.33%
MATCH: Indirect migrants									
Yes	55.29%	73.33%	56.59%	62.20%	46.87%	66.32%	62.84%	46.87%	57.66%
No	44.71%	26.67%	43.41%	37.80%	53.13%	33.68%	37.16%	53.13%	42.34%
Observations	1499	205	2011	1212	4920	5932	10859	4920	15779
Percentage	9.50%	1.30%	12.74%	7.68%	31.18%	37.59%	68.82%	31.18%	100%

Further, $MATCH_{i,t}$ is regressed sequentially on different control variables. As before we conduct the likelihood-ratio test for each set of control variables. The corresponding regression output and likelihood ratios are presented in Table 3.9. Table 3.16 presented in Appendix IV includes all the coefficients of control variables. The results of likelihood-ratio tests suggest that among the control variables only the dummy variable indicating gender of the individual should not be added to the regression equation. Our results reveal that higher education is positively associated with the probability of a good match. At the same time we can see that workers in smaller firms are more likely to perform a job corresponding to their initial training, whereas workers in larger firms are more frequently mismatched. Furthermore, jobs obtained after a long break are often associated with mismatch.

Table 3.9: Estimation results of occupational mismatch.

Variables	Dependent variable: MATCH								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EDU	0.056***	0.056***	0.053***	0.050***	0.051***	0.051***	0.051***	0.027***	0.027***
	(33.32)	(33.32)	(30.18)	(28.25)	(28.40)	(28.04)	(27.89)	(12.50)	(12.50)
AGE		-0.0061***	-0.0063***	-0.0057***	-0.0058***	-0.0057***	-0.0059***	-0.0058***	-0.0058***
		(-15.85)	(-16.04)	(-13.90)	(-14.07)	(-13.90)	(-14.15)	(-13.95)	(-13.95)
IND			v	v	v	v	v	v	v
TOJCH(Reference: First job)									
Job After Break				-0.106***	-0.108***	-0.108***	-0.106***	-0.090***	-0.090***
				(-4.73)	(-4.82)	(-4.84)	(-4.70)	(-3.93)	(-3.92)
Job With New Employer				-0.043*	-0.043*	-0.047*	-0.049*	-0.041	-0.041
				(-2.00)	(-2.01)	(-2.17)	(-2.27)	(-1.86)	(-1.86)
Company Taken Over				0.158***	0.160***	0.159***	0.158***	0.168***	0.168***
				(5.61)	(5.70)	(5.68)	(5.63)	(5.95)	(5.94)
Changed Job, Same Firm				0.037	0.046	0.044	0.045	0.035	0.035
				(1.52)	(1.85)	(1.79)	(1.82)	(1.36)	(1.35)
FSIZE(Reference: GE 2000)									
[1] LT 20					0.038**	0.042**	0.045***	0.087***	0.087***
					(2.92)	(3.23)	(3.41)	(6.50)	(6.50)
[2] GE 20 LT 200					0.006	0.011	0.014	0.041**	0.041**
					(0.50)	(0.86)	(1.06)	(3.03)	(3.03)
[3] GE 200 LT 2000					0.002	0.005	0.006	0.017	0.017
					(0.14)	(0.37)	(0.43)	(1.14)	(1.13)
STATE						v	v	v	v
Survey year t							v	v	v
ISEI								0.0073***	0.0073***
								(21.08)	(21.08)
FEMALE									-0.0015
									(-0.16)
	(-30.21)	(-16.52)	(-12.42)	(-8.42)	(-8.60)	(-7.84)	(-8.08)	(-9.45)	(-9.27)
LR test(Prob> χ^2)		0.00	0.00	0.00	0.0026	0.0002	0.0521	0.00	0.8761
Observations	15779	15779	15779	15779	15779	15779	15779	15779	15779
Pseudo R^2	0.059	0.070	0.085	0.093	0.093	0.095	0.097	0.118	0.118

t statistics in parentheses. Marginal effects for factor levels is the discrete change from the base level.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In the next step, $MIG_{i,t}$ is added to the regression equation. The coefficients are presented in Table 3.10 and the marginal effects are contained in squared brackets. Column (2) indicates that the coefficient on $MIG_{i,t}$ is negative and statistically significant, this means that foreign citizens are more likely to be mismatched inline with the descriptive statistics. The corresponding marginal effect reveals that foreign workers have 10% lower probability of being well matched in the job. This empirical evidence confirms our theoretical prediction from section 3.3. However, it is not only this negative link between being a foreigner and the probability of a good match that we want to test, but the underlying mechanism of the model based on the search channel. So we continue and add variable $REF_{i,t}$ to the regression equation in column (3). The coefficient of $REF_{i,t}$ is negative and statistically significant. This indicates that workers hired through referrals are more likely to be mismatched compared to those who are hired through the formal channel. Thus our empirical data confirms the model by Bentolila et al. (2010) and our assumption underlying the theoretical model in section 3.3.

Note, that after adding $REF_{i,t}$ to the regression equation the coefficient on $MIG_{i,t}$ becomes smaller in absolute value and the marginal effect of this variable is reduced from 10% down to 9.3%. Intuitively, this means the following. The fact that foreign workers rely more often on referral hiring explains a part of the negative link (0.7%) between being a foreigner and the probability of a good match. This result confirms the mechanism described by our theoretical model. However, the coefficient on $MIG_{i,t}$ stays negative and statistically significant after adding $REF_{i,t}$. This indicates that there are also other important reasons for the higher probability of mismatch in the group of foreign workers going beyond the search channel and not covered by our model.

Next, we empirically check if the two search channels exhibit different efficiency rates when used by different worker groups. Efficiency here refers to the probability of a good match. We do so by adding an interaction term $MIG_{i,t} \times REF_{i,t}$ into the regression, see column (4). The likelihood-ratio test suggests that $MIG_{i,t} \times REF_{i,t}$ should not be included into the regression equation since this variable is not significant. This means that referrals have equally low efficiency in generating good matches irrespective of the applicant's ethnic belonging.

When $CHAN_{i,t}$ is added to the regression equation instead of $REF_{i,t}$, the results are the following (see column (5)). The coefficients on $REF_{i,t}$ and $MIG_{i,t}$ are again negative and statistically significant. As in the descriptive statistics, referrals lead most often to mismatch compared to other search channels. Other search channels which are positively associated with mismatch are newspapers and the public employment agency. When we use detailed information about the search channel we can see that the marginal effect of variable MIG is reduced even further from 9.3% down to 9%. This means the following. The fact that foreign workers rely more often on newspapers and the public employment agency explains another 0.3% difference in the probability of mismatch between native and foreign workers. In specification (6) we additionally include the interaction terms between $MIG_{i,t}$ and $CHAN_{i,t}$, but none of these interaction terms is statistically significant. Moreover, the likelihood-ratio test suggests that the interaction terms should not be included to the regression equation. Again this shows that different search channels have similar match qualities when used by native and foreign workers. It is rather so that foreign workers are more likely to rely on search channels with lower efficiency, like referral hiring and employment agency.

Table 3.10: Estimation results of occupational mismatch by citizenship and search channels.

Variables	Dependent variable: MATCH					
	(1)	(2)	(3)	(4)	(5)	(6)
MIG		-0.099***	-0.093***	-0.092***	-0.090***	0.003
citizen		(-5.54)	(-5.14)	(-4.00)	(-5.00)	(0.04)
REF			-0.103***	-0.103***		
			(-10.87)	(-10.47)		
MIG × REF				-0.002		
				(-0.05)		
CHAN (Reference: Internet)						
Public emp. agency					-0.078***	-0.076***
					(-3.70)	(-3.49)
Private emp. agency					-0.062	-0.060
					(-1.53)	(-1.44)
Newspaper					-0.056**	-0.056**
					(-2.83)	(-2.74)
Referrals					-0.129***	-0.125***
					(-7.33)	(-6.93)
Other					0.004	0.010
					(0.23)	(0.55)
MIG × CHAN(Reference: MIG × Internet)						
MIG × Public empl. agency						-0.073
						(-0.73)
MIG × Private empl. agency						-0.066
						(-0.41)
MIG × Newspaper						-0.048
						(-0.50)
MIG × Referrals						-0.098
						(-1.11)
MIG × Other						-0.137
						(-1.52)
Control variables	v	v	v	v	v	v
Time FE	v	v	v	v	v	v
LR test(Prob> χ^2)		0.00	0.00	0.9612	0.00	0.5276
Observations	15779	15779	15779	15779	15779	15779
Pseudo R^2	0.118	0.119	0.125	0.125	0.127	0.127

t statistics in parentheses. Marginal effects for factor levels is the discrete change from the base level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Control variables in Table 3.10 include age, education, industry, Standard International Socio-Economic Index of Occupational Status of individuals, firm size with 4 categories, state of residence, survey year, and the type of job change. Table 3.17 presented in Appendix V includes all the coefficients of control variables.

In Table 3.11 we substitute binary variable $MIG_{i,t}$ with a more detailed variable $MIGBACK_{i,t}$ containing three categories. Column (2) shows that the coefficient for direct migrants is negative and statistically significant, while the coefficient for indirect migrants is not statistically significant. This means that compared to German nationals direct migrants are less likely to be well matched, while indirect migrants can not be statistically distinguished from native German workers. The marginal effect shows that direct migrants are 8.7% more likely to be mismatched than German nationals. Next, $REF_{i,t}$ is added to the regression in column (3). The negative and statistically significant coefficient of referrals suggests that referral hiring leads to good matches less often compared to hiring through formal search channels. We can see that the marginal effect is again reduced from 8.7% down to 8%. This confirms our earlier conclusion that 0.7% of the differences in mismatch rates between migrant and native workers is due to the fact that migrants rely more often on their social networks. Now we can additionally conclude that this effect is largely generated by direct migrants. The interaction terms in column (4) are again insignificant.

Further, we include a more detailed variable $CHAN_{i,t}$ instead of a binary indicator $REF_{i,t}$ for the search channel. The marginal effect of being a direct migrant falls from 8% down to 7.7%, so this regression confirms the fact that additional 0.3% difference in the probabilities of mismatch is due to the fact that direct migrants use less efficient search channels such as newspapers and services of the public employment agency more often than native German workers. At this step we decided not to include the interaction terms between the search channels and $MIGBACK_{i,t}$ as none of the interaction terms was significant in the previous regressions.

Table 3.11: Estimation results of occupational mismatch using migration background and search channels.

Variables	Dependent variable: MATCH				
	(1)	(2)	(3)	(4)	(5)
MIGBACK (Reference: German national)					
Direct migrant		-0.087***	-0.080***	-0.079***	-0.077***
		(-5.81)	(-5.33)	(-4.19)	(-5.15)
Indirect migrant		-0.029	-0.027	-0.030	-0.026
		(-1.65)	(-1.50)	(-1.37)	(-1.44)
Referrals			-0.103***	-0.103***	
			(-10.81)	(-9.89)	
MIGBACK × REF (Reference: German national × Formal channels)					
Direct migrant × referrals				-0.002	
				(-0.06)	
Indirect migrant × referrals				0.009	
				(0.25)	
Chan (Reference: Internet)					
Public emp. agency					-0.079***
					(-3.75)
Private emp. agency					-0.061
					(-1.51)
Newspaper					-0.058**
					(-2.92)
Referrals					-0.130***
					(-7.39)
Other					0.002
					(0.12)
Control variables	v	v	v	v	v
Time FE	v	v	v	v	v
LR test(Prob> χ^2)		0.00	0.00	0.9666	0.00
Observations	15779	15779	15779	15779	15779
Pseudo R^2	0.118	0.119	0.125	0.125	0.127

Continued on next page

Variables	Dependent variable: MATCH				
	(1)	(2)	(3)	(4)	(5)
t statistics in parentheses. Marginal effects for factor levels is the discrete change from the base level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$					

Control variables in Table 3.11 include age, education, industry, Standard International Socio-Economic Index of Occupational Status of individuals, firm size with 4 categories, state of residence, survey year, and the type of job change. Table 3.18 presented in Appendix VI includes all the coefficients of control variables.

In addition to the estimations described above, we exploit individual effects estimations using the panel nature of the dataset. We consider individual effects estimations, because there might be unobservable factors that correlate both with the use of social networks and working mismatched. Appendix VII summarizes empirical approach which is used to deal with issues which may be caused by the potential endogenous variables in the regressions. Panel probit regression model with random effects is used to check the robustness of the results obtained above using logistic regression model. The results of panel probit regression with random effects confirm all the conclusions derived from the results of logistic regression. The results are qualitatively similar to the results, where logistic regression model is used. But the magnitude of the effects of variables standing for migration background and search channels is different (see Appendix VIII, IX and X).

3.4.2 Blinder-Oaxaca Decomposition

In the previous subsection we found that search channels have a significant effect on the probability of being well matched. Also, the probability of being well matched is different for German and foreign citizens. The goal of this section is to quantify how much of the difference in mismatch rates between the two groups can be explained by differences in the search channels. For this purpose we use the Blinder-Oaxaca decomposition applied to the linear probability model of the outcome variable $MATCH_{i,t}$[3]. This decomposition is based on the following equation:

$$\hat{Y}_N - \hat{Y}_F = \underbrace{(\bar{X}_N - \bar{X}_F)'\hat{B}_N}_{\text{Endowment effect (explained)}} + \underbrace{\bar{X}_F'(\hat{B}_N - \hat{B}_F)}_{\text{Coefficient effect (unexplained)}} \tag{3.2}$$

Here $\hat{Y}_N$ is the estimated proportion of well matched German citizens, and $\hat{Y}_F$ is the estimated proportion of well matched foreign citizens. $\bar{X}_N$ and $\bar{X}_F$ are the vectors of average characteristics (endowments) of German and foreign citizens respectively. $\hat{B}_N$ and $\hat{B}_F$ are the estimated coefficient vectors for the two groups. Note, that in the above-mentioned two-fold decomposition the coefficients of the majority group are assumed to be nondiscriminatory. The first element on the right-hand side shows differences in the proportions

[3]Blinder-Oaxaca decomposition was conducted using "oaxaca" command in the statistical program Stata. See details at Jann et al. (2008).

of well matched workers stemming from different endowments of the two worker groups. This includes observable individual characteristics, such as education, gender and age, but also the search channel. The second element on the right-hand side shows remaining differences in the proportions of well-matched workers which can not be explained by the regression.

The estimation results of the Blinder-Oaxaca decomposition are presented in Table 3.12.

Table 3.12: Estimation results of Blinder-Oaxaca decomposition by citizenship.

	Coefficient	Std.Err.		Coefficient	Std.Err.
German citizens	0.5756***	0.0041	German citizens	14754	
Foreign citizens	0.4205***	0.0157	Foreign citizens	1022	
Difference	0.1551***	0.0163	Observations	15776	
Endowment effect	0.0756***	0.0067	Coefficient effect	0.0796***	0.0151
Public emp. ag.	0.00031	0.00028	Public emp. ag.	0.00028	0.00490
Private emp. ag.	0.00002	0.00011	Private emp. ag.	0.00027	0.00164
Newspapers	0.00004	0.00017	Newspapers	-0.00318	0.00588
Internet	0.00143**	0.00053	Internet	-0.00349	0.00296
Referrals	0.00760***	0.00146	Referrals	0.00535	0.01370
Other	0.00614***	0.00128	Other	0.01801	0.00982
Firm size	-0.00428	0.00122	Firm size	-0.00470	0.00644
Industry	0.00984***	0.00258	Industry	0.07943	0.05027
TOJCH	0.00336	0.00152	TOJCH	0.04663*	0.02383
State	-0.00566*	0.00257	State	-0.04982	0.03364
Time	0.00017	0.00119	Time	-0.00308	0.00690
Education	0.02558***	0.00293	Education	-0.02489	0.08296
Age	-0.00592***	0.00177	Age	-0.08710	0.05564
ISEI	0.03697***	0.00354	ISEI	-0.08702*	0.04400

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The estimated fraction of well matched German citizens is 57.56%, and the estimated fraction of well matched foreign citizens is 42.05%. So the difference is equal to 15.51%. We already know these numbers from Table 3.7. The decomposition shows that the endowment effect is equal to 7.56%, while the unexplained coefficient effect is 7.96%. Thus our regression can explain roughly a half of the observable difference in the mismatch rates between foreign and German workers.

We can see that two variables that explain the largest part of the gap in mismatch rates are education and the ISEI index of occupational prestige. This means that foreign workers in Germany are less educated on average, and overrepresented in low skill jobs with low occupational prestige. At the same time these jobs are associated with higher mismatch probability compared to high skill jobs with high occupational prestige. Combined together these effects explain $2.6\% + 3.7\% = 6.3\%$ out of the endowment effect equal to 7.6%. This effect is reduced by 0.6% because German workers are older on average and the probability of mismatch is increasing with age. At the same time foreign workers are overrepresented in industries with higher occupational mismatch (such as transportation and trade), which explains another 1% of the endowment effect. So the part of the endowment effect which is due to differences in the industry and observable worker characteristics can be estimated as $6.7/7.6 = 88.2\%$. Finally, Table 3.12 shows that additional 0.9% of the endowment effect are explained by the fact that foreign workers use less efficient search channels compared to German workers. So the part of the endowment effect which is due to the different search channels can be estimated as $0.9/7.6 = 11.8\%$. Note that most of this effect is because of the more intensive referral hiring in the group of foreign workers (0.76% out of 0.9) with only a small contribution of the internet (0.14 out of 0.9) and insignificant contributions of other channels, such as the public employment agency. Even though the public employment agency produces substantial mismatch comparable to that of referral hiring (see Table 3.8) there are no large differences in the utilization of this channel between native and immigrant workers (see Table 3.7), so it doesn't contribute to the gap in mismatch rates between these two groups. This supports our theoretical approach where we isolate referral hiring from all other (formal) channels and treat it separately.

To conclude, first, both the estimations and the Blinder-Oaxaca decomposition results show that there is significant difference in the proportions of good matches between German citizens and foreign citizens equal to 15.1%. Second, those who are matched through referrals are more likely to be mismatched compared to those who are matched through formal channels. Moreover, the results of the Blinder-Oaxaca decomposition show that explanatory variables used in the estimation account for about a half of the total gap in mismatch rates, which is the endowment effect. And finally, the fact the foreign workers use less efficient search channels, such as referral hiring, account for 11.8% of the endowment effect with the remaining gap attributed to education, occupational prestige, age and industry differences.

3.5 Conclusions

In this study we investigate the link between the job search channels and occupational mismatch with a specific focus on differences between native and immigrant workers. We use data from the German Socio-Economic Panel (SOEP) over the period 2000-2014. First, we find that referral hiring via social networks is the most frequent single channel of generating jobs in Germany. Moreover, this channel is used more frequently by immigrant workers rather than natives. This could be due to the higher risk of unemployment that immigrant workers are confronted with and larger difficulties of finding jobs in a formal way. In this case social networks and referral hiring serve as a channel of last resort for the immigrant population.

We combine this empirical evidence with the finding by Bentolila et al. (2010) that referral hiring generates more occupational mismatch than formal search. The reason is that workers tend to send formal applications to jobs in their primary occupation, whereas friends and relatives providing job recommendations often work in different occupations giving rise to occupational mismatch. We incorporate this empirical evidence into a search and matching model with two ethnic worker groups (natives and immigrants), two occupations and two search channels (formal applications and informal network hiring). Job recommendations are given by employed workers to the unemployed friends in their social network. We assume that all workers have the same size of social networks, but their composition differs across groups. In particular, we take into account that social networks exhibit ethnic and professional homophily meaning biased link formation towards friends with the same ethnicity and from the same profession. Our model predicts that more intensive utilisation of referal hiring leads to more frequent occupational mismatch of immigrant workers. One condition for this result is that the gap in the job destruction rates between native and immigrant workers is not too large which is satisfied for a realistic parameter setting motivated by the data. From a theoretical perspective this result strongly depends on the degree of professional homophily characterising social networks but it is not sensitive to the network size.

Next we test the underlying assumption of the model and find empirical support for the fact that referral hiring generates more occupational mismatch than formal search. The data reveals that referral hiring is the least efficient job creating channel in terms of match quality among public and private employment agencies, specialised newspapers, direct applications in internet and other channels. Further, we test the theoretical prediction of our model that differences in the incidence of referral hiring between native and immigrant workers contribute significantly to the gap in mismatch rates between these groups. To achieve this goal we perform a Blinder-Oaxaca decomposition. The overall gap in the mismatch rates is equal to 15.5%. Roughly a half of this effect (7.6%) can be explained by observable differences in the endowments between native and immigrant workers including the search channel. We find that differences in the search strategies explain about 1% of the gap in the mismatch rates. This effect is significant with the remaining gap (6.6%) attributed to education, age and industry differences. This confirms our theoretical prediction that at least a part of the mismatch gap between native and immigrant workers is due to the less efficient job search channels used by immigrant workers. Finally, our decomposition reveals that only referral hiring and internet contribute significantly to the gap in mismatch rates but not other channels, such as the public employment agency. It is despite the fact that the public employment agency generates high rates of occupational mismatch comparable to referrals. This finding supports our theoretical approach where we reduced the number of search channels to referral hiring and one formal channel.

3.6 Acknowledgement

We would like to thank the members of the ZIF Research Group "In Search of the Global Labour Market: Actors, Structures and Policies" for the continuing support and contributions. This paper also benefited from the useful comments by Herbert Dawid and Michele Battisti. Anna Zaharieva acknowledges financial support of the Center for Interdisciplinary Research (ZIF) at Bielefeld University. Sevak Alaverdyan acknowledges financial support

of the German Academic Exchange Service (DAAD).

3.7 Appendix

Proof of lemma 1: Without ethnic homophily we know that $n_{jN}^{AA} = (1-h)\gamma n$ and $n_{jF}^{AA} = h\gamma n$, $j = N, F$. So variables M_{NN}^{AA} and M_{FN}^{AA} can be reduced to:

$$M_{jN}^{AA} = sv^A \frac{m_j^A}{m_N^A + m_F^A}\left(1 - \left[\frac{e_N^A}{1-h}\right]^{(1-h)\gamma n}\left[\frac{e_F^A}{h}\right]^{h\gamma n}\right)\frac{u_N^A}{u_N^A + u_F^A}$$

$$\begin{aligned}\mu_N^{AA} &= \frac{M_{NN}^{AA} + M_{FN}^{AA}}{u_N^A} = \frac{sv^A}{u_N^A}\left(\frac{m_N^A + m_F^A}{m_N^A + m_F^A}\right)\left(1 - \left[\frac{e_N^A}{1-h}\right]^{(1-h)\gamma n}\left[\frac{e_F^A}{h}\right]^{h\gamma n}\right)\frac{u_N^A}{u_N^A + u_F^A} \\ &= \frac{sv^A}{u_N^A + u_F^A}\left(1 - \left[\frac{e_N^A}{1-h}\right]^{(1-h)\gamma n}\left[\frac{e_F^A}{h}\right]^{h\gamma n}\right)\end{aligned}$$

Further we can also rewrite variables M_{NF}^{AA} and M_{FF}^{AA} as:

$$M_{jF}^{AA} = sv^A \frac{m_j^A}{m_N^A + m_F^A}\left(1 - \left[\frac{e_N^A}{1-h}\right]^{(1-h)\gamma n}\left[\frac{e_F^A}{h}\right]^{h\gamma n}\right)\frac{u_F^A}{u_N^A + u_F^A}$$

$$\begin{aligned}\mu_F^{AA} &= \frac{M_{NF}^{AA} + M_{FF}^{AA}}{u_F^A} = \frac{sv^A}{u_F^A}\left(\frac{m_N^A + m_F^A}{m_N^A + m_F^A}\right)\left(1 - \left[\frac{e_N^A}{1-h}\right]^{(1-h)\gamma n}\left[\frac{e_F^A}{h}\right]^{h\gamma n}\right)\frac{u_F^A}{u_N^A + u_F^A} \\ &= \frac{sv^A}{u_N^A + u_F^A}\left(1 - \left[\frac{e_N^A}{1-h}\right]^{(1-h)\gamma n}\left[\frac{e_F^A}{h}\right]^{h\gamma n}\right)\end{aligned}$$

so that $\mu_N^{AA} = \mu_F^{AA}$.

Appendix I. Estimation results: employment rates.

Table 3.13: Estimation results of employment rates, full table.

Variables	Dependent variable: EMP							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EDU	0.020***	0.019***	0.019***	0.019***	0.019***	0.019***	0.019***	0.018***
	(94.84)	(94.76)	(94.93)	(94.98)	(101.42)	(101.70)	(100.00)	(90.46)
AGE		0.00014**	0.00014**	-0.00064***	-0.00049***	-0.00050***	-0.00088***	-0.00095***
		(3.24)	(3.25)	(-13.47)	(-11.28)	(-11.55)	(-18.46)	(-19.98)
FEMALE			-0.0086***	-0.0063***	-0.0058***	-0.0059***	-0.0063***	-0.0069***
			(-8.70)	(-6.58)	(-6.86)	(-6.97)	(-7.36)	(-8.08)
MARST(Reference: Married)								
[2] Single				-0.058***	-0.045***	-0.045***	-0.058***	-0.061***
				(-33.51)	(-28.97)	(-29.03)	(-31.92)	(-33.02)
[3] Widowed				-0.017***	-0.014***	-0.014***	-0.016***	-0.017***

Continued on next page

Variables	Dependent variable: EMP							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				(-4.90)	(-4.54)	(-4.51)	(-5.13)	(-5.42)
[4] Divorced				-0.050***	-0.043***	-0.043***	-0.045***	-0.047***
				(-25.75)	(-24.47)	(-24.52)	(-25.78)	(-26.46)
[5] Separated				-0.048***	-0.043***	-0.043***	-0.045***	-0.045***
				(-12.87)	(-12.53)	(-12.52)	(-13.09)	(-13.20)
Number of							-0.009***	-0.009***
Children in HH							(-20.08)	(-19.07)
Foreign								-0.043***
citizen								(-19.57)
STATE (Reference: Schleswig-Holstein)								
[2] Hamburg					0.003	0.003	0.002	0.007
					(0.64)	(0.64)	(0.55)	(1.50)
[3] Lower Saxony					0.010***	0.010***	0.010***	0.011***
					(3.57)	(3.56)	(3.66)	(3.82)
[4] Bremen					-0.028***	-0.028***	-0.029***	-0.029***
					(-4.15)	(-4.15)	(-4.29)	(-4.12)
[5] North-Rhine					0.002	0.002	0.002	0.005
-Westfalia					(0.85)	(0.86)	(0.75)	(1.92)
[6] Hessen					0.013***	0.013***	0.012***	0.016***
					(4.44)	(4.46)	(4.30)	(5.50)
[7] Rheinland-Pfalz					0.015***	0.015***	0.015***	0.017***
					(5.17)	(5.20)	(5.25)	(5.61)
[8] Baden					0.024***	0.024***	0.024***	0.029***
-Wuerttemberg					(9.34)	(9.34)	(9.37)	(10.90)
[9] Bavaria					0.025***	0.025***	0.025***	0.027***
					(9.84)	(9.84)	(9.71)	(10.20)
[10] Saarland					0.012*	0.012*	0.011**	0.013**
					(2.97)	(2.97)	(2.59)	(3.05)
[11] Berlin					-0.042***	-0.042***	-0.042***	-0.039***
					(-10.35)	(-10.33)	(-10.36)	(-9.54)
[12] Brandenburg					-0.072***	-0.072***	-0.073***	-0.077***
					(-17.03)	(-17.02)	(-17.19)	(-17.51)
[13] Mecklenburg					-0.067***	-0.067***	-0.067***	-0.073***
-Vorpommern					(-13.32)	(-13.30)	(-13.47)	(-13.82)
[14] Saxony					-0.049***	-0.049***	-0.051***	-0.054***
					(-14.34)	(-14.34)	(-14.79)	(-15.25)
[15] Saxony-Anhalt					-0.076***	-0.076***	-0.078***	-0.083***
					(-17.79)	(-17.81)	(-18.15)	(-18.56)
[16] Thuringia					-0.053***	-0.053***	-0.055***	-0.059***
					(-13.40)	(-13.40)	(-13.80)	(-14.24)
Survey year t (Reference: 2000)								
2001						-0.00013	0.00001	-0.00003
						(-0.06)	(0.01)	(-0.02)
2002						-0.00144	-0.00135	-0.00167
						(-0.65)	(-0.59)	(-0.73)
2003						-0.00788***	-0.00784**	-0.00825***
						(-3.37)	(-3.27)	(-3.44)
2004						-0.01044***	-0.01050***	-0.01092***
						(-4.37)	(-4.29)	(-4.45)
2005						-0.00926***	-0.00942***	-0.00974***
						(-3.83)	(-3.80)	(-3.93)
2006						-0.01004***	-0.01020***	-0.01089***
						(-4.20)	(-4.16)	(-4.42)
2007						0.00009	0.00015	-0.00050
						(0.04)	(0.06)	(-0.21)

Continued on next page

Variables	Dependent variable: EMP							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2008						0.00602**	0.00628**	0.00565*
						(2.67)	(2.72)	(2.44)
2009						0.00382	0.00419	0.00348
						(1.70)	(1.82)	(1.50)
2010						-0.00768***	-0.00387	-0.00486*
						(-3.44)	(-1.74)	(-2.17)
2011						0.00064	0.00431*	0.00332
						(0.30)	(2.05)	(1.57)
2012						0.00109	0.00477*	0.00386
						(0.52)	(2.28)	(1.83)
2013						-0.00402	-0.00040	0.00172
						(-1.95)	(-0.19)	(0.84)
2014						-0.00078	0.00272	0.00420*
						(-0.37)	(1.28)	(2.01)
LR test(Prob> χ^2)		0.0012	0.00	0.00	0.00	0.00	0.00	0.00
Observations	213592	213592	213592	213592	213592	213592	213592	213592
Pseudo R^2	0.062	0.062	0.063	0.081	0.116	0.117	0.120	0.125

t statistics in parentheses. Marginal effects for factor levels is the discrete change from the base level.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix II. Estimation results: referral hiring.

Table 3.14: Estimation results of referral hiring, full table.

Variables	Dependent variable: REF									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
EDU	-0.023***	-0.023***	-0.021***	-0.017***	-0.015***	-0.014***	-0.014***	-0.014***	-0.006***	-0.006***
	(-17.59)	(-17.44)	(-16.13)	(-12.62)	(-10.83)	(-9.69)	(-9.64)	(-9.76)	(-3.63)	(-3.62)
AGE		-0.00063*	-0.00062*	-0.00094**	-0.00090**	-0.00057	-0.00056	-0.00049	-0.00062	-0.00063
		(-2.11)	(-2.08)	(-3.13)	(-2.97)	(-1.74)	(-1.72)	(-1.48)	(-1.88)	(-1.89)
Foreign			0.085***	0.085***	0.082***	0.077***	0.075***	0.077***	0.074***	0.072***
citizen			(6.41)	(6.33)	(6.13)	(5.80)	(5.54)	(5.75)	(5.53)	(5.44)
FSIZE(Reference: GE 2000)										
[1] LT 20				0.154***	0.152***	0.110***	0.110***	0.108***	0.097***	0.098***
				(15.99)	(15.54)	(10.81)	(10.79)	(10.60)	(9.33)	(9.39)
[2] GE 20 LT 200				0.077***	0.076***	0.039***	0.040***	0.037***	0.030**	0.030**
				(8.03)	(7.82)	(3.89)	(3.94)	(3.69)	(2.93)	(2.91)
[3] GE 200 LT 2000				0.028**	0.031**	0.008	0.008	0.007	0.004	0.004
				(2.67)	(2.89)	(0.71)	(0.73)	(0.61)	(0.36)	(0.33)
IND(Reference: Services)										
[1] Agriculture					-0.003	-0.003	0.001	0.0005	-0.021	-0.026
					(-0.10)	(-0.10)	(0.03)	(0.02)	(-0.82)	(-1.03)
[2] Energy					-0.070	-0.058	-0.058	-0.053	-0.050	-0.054
					(-1.94)	(-1.57)	(-1.57)	(-1.42)	(-1.31)	(-1.42)
[3] Mining					0.095	0.125	0.122	0.116	0.110	0.102
					(0.97)	(1.22)	(1.19)	(1.13)	(1.08)	(1.00)
[4] Manufacturing					0.033**	0.025*	0.025*	0.023*	0.021*	0.016
					(3.08)	(2.40)	(2.37)	(2.21)	(1.99)	(1.46)
[5] Construction					0.004	-0.003	-0.001	-0.003	-0.008	-0.017
					(0.33)	(-0.25)	(-0.09)	(-0.30)	(-0.78)	(-1.45)
[6] Trade					0.050***	0.040***	0.040***	0.039***	0.041***	0.040***
					(5.12)	(4.13)	(4.16)	(4.05)	(4.25)	(4.17)

Continued on next page

Variables	Dependent variable: REF									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
[7] Transport					0.037*	0.029	0.030	0.027	0.023	0.017
					(2.34)	(1.89)	(1.94)	(1.75)	(1.51)	(1.10)
[8] Bank,Insurance					-0.041*	-0.020	-0.020	-0.022	-0.009	-0.011
					(-2.12)	(-0.95)	(-0.96)	(-1.09)	(-0.43)	(-0.51)
TOJCH(Reference: First job)										
Job After Break						-0.060***	-0.060***	-0.068***	-0.074***	-0.072***
						(-4.58)	(-4.57)	(-5.18)	(-5.60)	(-5.49)
Job With New Employer						0.030*	0.030*	0.034**	0.031*	0.031*
						(2.42)	(2.40)	(2.77)	(2.52)	(2.51)
Company Taken Over						-0.243***	-0.242***	-0.243***	-0.246***	-0.246***
						(-14.25)	(-14.23)	(-14.26)	(-14.49)	(-14.54)
Changed Job, Same Firm						-0.255***	-0.255***	-0.256***	-0.256***	-0.256***
						(-18.41)	(-18.41)	(-18.48)	(-18.36)	(-18.36)
STATE(Reference: Schleswig-Holstein)										
[2] Hamburg							-0.0012			
							(-0.04)			
[3] Lower Saxony							-0.0182			
							(-0.79)			
[4] Bremen							-0.0144			
							(-0.34)			
[5] North-Rhine -Westfalia							-0.0002			
							(-0.01)			
[6] Hessen							-0.0059			
							(-0.25)			
[7] Rheinland-Pfalz							-0.0107			
							(-0.42)			
[8] Baden-Wuerttemberg							-0.0116			
							(-0.52)			
[9] Bavaria							-0.0057			
							(-0.26)			
[10] Saarland							0.0069			
							(0.18)			
[11] Berlin							-0.0010			
							(-0.04)			
[12]Brandenburg							-0.0113			
							(-0.44)			
[13]Mecklenburg -Vorpommern							-0.0639*			
							(-2.26)			
[14] Saxony							-0.0231			
							(-0.97)			
[15] Saxony-Anhalt							-0.0051			
							(-0.20)			
[16] Thuringia							0.0155			
							(0.59)			
Survey year t (Reference: 2000)										
2001								-0.0034	-0.0028	-0.0028
								(-0.22)	(-0.18)	(-0.18)
2002								-0.0158	-0.0150	-0.0146
								(-0.98)	(-0.92)	(-0.90)
2003								0.0192	0.0206	0.0207
								(1.11)	(1.19)	(1.19)
2004								-0.0039	-0.0037	-0.0042
								(-0.23)	(-0.22)	(-0.25)
2005								0.0136	0.0140	0.0134
								(0.78)	(0.79)	(0.76)

Continued on next page

Variables	Dependent variable: REF									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
2006								0.0020	0.0009	0.0004
								(0.12)	(0.05)	(0.02)
2007								0.0461**	0.0451**	0.0450**
								(2.69)	(2.63)	(2.63)
2008								0.0326	0.0315	0.0315
								(1.91)	(1.84)	(1.84)
2009								0.0348*	0.0354*	0.0351*
								(2.04)	(2.08)	(2.06)
2010								-0.0440**	-0.0452**	-0.0448**
								(-2.62)	(-2.69)	(-2.67)
2011								-0.0308	-0.0314	-0.0309
								(-1.75)	(-1.78)	(-1.75)
2012								-0.0246	-0.0272	-0.0267
								(-1.34)	(-1.48)	(-1.45)
2013								-0.0333	-0.0325	-0.0319
								(-1.83)	(-1.79)	(-1.75)
2014								-0.0111	-0.0142	-0.0124
								(-0.69)	(-0.89)	(-0.78)
ISEI									-0.0022***	-0.0022***
									(-7.95)	(-8.00)
FEMALE										-0.0157*
										(-2.10)
LR test(Prob> χ^2)		0.0344	0.00	0.0275	0.00	0.00	0.5708	0.00	0.00	0.00
Observations	19148	19148	19148	19148	19148	19148	19148	19148	19148	19148
Pseudo R^2	0.013	0.014	0.015	0.028	0.030	0.058	0.058	0.060	0.062	0.062

t statistics in parentheses. Marginal effects for factor levels is the discrete change from the base level.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix III. Descriptive statistics of variables used as control variables.

Table 3.15: Descriptive statistics of control variables.

	German citizens	Foreign citizens	German nationals	Direct migrants	Indirect migrants	Overall
EDU	12.92	11.65	12.96	12.01	12.57	12.84
AGE	35.15	36.26	36.42	37.33	31.59	36.19
IND						
Agriculture	1.54%	1.85%	1.61%	1.53%	0.97%	1.56%
Energy	1.00%	0.29%	0.99%	0.58%	1.07%	0.95%
Mining	0.13%	0.20%	0.14%	0.13%	0.00%	0.13%
Manufacturing	13.98%	17.66%	13.66%	17.84%	15.79%	14.22%
Construction	11.89%	13.07%	11.86%	12.72%	12.11%	11.97%
Trade	17.43%	22.24%	17.33%	20.52%	18.80%	17.75%
Transport	5.68%	8.98%	5.49%	7.86%	8.04%	5.89%
Bank,Insurance	3.46%	1.85%	3.61%	1.41%	3.10%	3.36%
Services	44.90%	33.85%	45.30%	37.40%	40.12%	44.18%
TOJCH						
First job	5.14%	4.98%	5.11%	4.54%	6.30%	5.13%
Job After Break	26.94%	29.66%	27.02%	29.92%	24.03%	27.11%
Job With New Employer	54.58%	54.05%	54.45%	57.67%	54.94%	54.80%
Company Taken Over	3.33%	3.12%	3.27%	2.81%	4.75%	3.32%
Changed Job, Same Firm	10.01%	4.20%	10.15%	5.05%	9.98%	9.63%

Continued on next page

	German citizens	Foreign citizens	German nationals	Direct migrants	Indirect migrants	Overall
FSIZE						
[1] LT 20	33.12%	38.44%	33.24%	36.13%	32.36%	33.47%
[2] GE 20 LT 200	29.77%	29.56%	29.65%	31.84%	28.00%	29.76%
[3] GE 200 LT2000	17.86%	16.00%	17.86%	16.11%	18.70%	17.74%
[4] GE 2000	19.24%	16.00%	19.25%	15.92%	20.93%	19.03%
STATE						
[1] Schleswig-Holstein	2.95%	1.17%	3.01%	2.11%	1.65%	2.83%
[2] Hamburg	1.76%	0.68%	1.71%	1.02%	2.42%	1.69%
[3] Lower Saxony	9.36%	7.02%	9.33%	10.36%	5.91%	9.21%
[4] Bremen	0.83%	0.78%	0.74%	1.53%	0.78%	0.82%
[5] North-Rhine-Westfalia	18.34%	24.29%	17.79%	24.68%	21.71%	18.73%
[6] Hessen	7.45%	8.98%	7.15%	9.08%	10.27%	7.55%
[7] Rheinland-Pfalz	4.28%	4.78%	4.04%	6.01%	5.33%	4.32%
[8] Baden-Wuerttemberg	10.89%	23.41%	10.04%	18.73%	22.38%	11.71%
[9] Bavaria	14.38%	20.49%	14.36%	16.88%	16.96%	14.78%
[10] Saarland	0.92%	1.85%	0.83%	2.24%	1.07%	0.98%
[11] Berlin	3.86%	4.39%	3.85%	4.35%	3.78%	3.89%
[12] Brandenburg	4.75%	0.78%	5.11%	1.15%	1.74%	4.49%
[13] Mecklenburg-Vorpommern	2.79%	0.20%	3.02%	0.19%	1.16%	2.62%
[14] Saxony	7.89%	0.88%	8.56%	1.02%	2.81%	7.43%
[15] Saxony-Anhalt	4.90%	0.00%	5.29%	0.32%	0.10%	4.58%
[16] Thuringia	4.66%	0.29%	5.19%	0.32%	0.10%	4.37%
Survey year t						
2000	9.65%	13.85%	9.89%	10.81%	9.01%	9.92%
2001	8.17%	10.63%	8.32%	8.76%	7.75%	8.33%
2002	6.84%	8.10%	7.06%	6.14%	6.30%	6.92%
2003	6.01%	6.24%	6.09%	4.86%	6.88%	6.02%
2004	6.36%	5.76%	6.44%	6.27%	4.94%	6.32%
2005	5.96%	4.98%	5.99%	4.80%	6.30%	5.89%
2006	6.66%	5.56%	6.71%	5.88%	6.10%	6.59%
2007	7.16%	6.44%	7.15%	7.23%	6.59%	7.12%
2008	7.00%	5.85%	6.91%	6.65%	7.56%	6.93%
2009	7.14%	7.22%	7.07%	7.93%	6.88%	7.14%
2010	5.91%	5.27%	5.86%	5.63%	6.40%	5.87%
2011	5.63%	3.22%	5.48%	4.86%	6.40%	5.48%
2012	4.98%	3.32%	4.96%	3.90%	5.23%	4.87%
2013	5.04%	3.22%	4.99%	3.90%	5.52%	4.92%
2014	7.49%	10.34%	7.08%	12.40%	8.14%	7.67%
ISEI	45.51	39.02	45.72	39.38	45.69	45.09
Gender						
Male	43.99%	54.34%	43.94%	46.68%	50.78%	44.66%
Female	56.01%	45.66%	56.06%	53.32%	49.22%	55.34%
Observations	14754	1025	13183	1564	1032	15779
Percentage	93.50%	6.50%	83.55%	9.91%	6.54%	100%

Appendix IV. Estimation results: occupational mismatch.

Table 3.16: Estimation results of occupational mismatch.

Variables	Dependent variable: MATCH								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EDU	0.056***	0.056***	0.053***	0.050***	0.051***	0.051***	0.051***	0.027***	0.027***

Continued on next page

Variables	Dependent variable: MATCH								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	(33.32)	(33.32)	(30.18)	(28.25)	(28.40)	(28.04)	(27.89)	(12.50)	(12.50)
AGE		-0.0061***	-0.0063***	-0.0057***	-0.0058***	-0.0057***	-0.0059***	-0.0058***	-0.0058***
		(-15.85)	(-16.04)	(-13.90)	(-14.07)	(-13.90)	(-14.15)	(-13.95)	(-13.95)
IND(Reference: Services)									
[1] Agriculture			-0.054	-0.043	-0.050	-0.037	-0.037	0.032	0.031
			(-1.62)	(-1.30)	(-1.49)	(-1.11)	(-1.09)	(0.97)	(0.95)
[2] Energy			-0.111**	-0.126**	-0.117**	-0.116**	-0.119**	-0.141**	-0.141**
			(-2.56)	(-2.89)	(-2.67)	(-2.63)	(-2.70)	(-3.19)	(-3.19)
[3] Mining			-0.229 *	-0.243*	-0.235*	-0.216	-0.215	-0.211	-0.212
			(-2.05)	(-2.17)	(-2.09)	(-1.89)	(-1.88)	(-1.81)	(-1.82)
[4] Manufacturing			-0.051***	-0.054***	-0.049***	-0.052***	-0.049***	-0.043**	-0.043**
			(-4.01)	(-4.20)	(-3.80)	(-4.02)	(-3.79)	(-3.28)	(-3.23)
[5] Construction			0.077***	0.078***	0.078***	0.082***	0.084***	0.099***	0.099***
			(6.12)	(6.19)	(6.16)	(6.46)	(6.62)	(7.87)	(7.34)
[6] Trade			-0.103***	-0.098***	-0.101***	-0.102***	-0.101***	-0.109***	-0.109***
			(-8.66)	(-8.16)	(-8.35)	(-8.47)	(-8.34)	(-8.92)	(-8.92)
[7] Transport			-0.212***	-0.218***	-0.215***	-0.216***	-0.214***	-0.202***	-0.202***
			(-11.79)	(-12.06)	(-11.83)	(-11.83)	(-11.72)	(-10.84)	(-10.70)
[8] Bank,Insurance			0.054*	0.039	0.045	0.044	0.043	-0.002	-0.003
			(2.37)	(1.66)	(1.92)	(1.69)	(1.80)	(-0.10)	(-0.10)
TOJCH(Reference: First job)									
Job After Break				-0.106***	-0.108***	-0.108***	-0.106***	-0.090***	-0.090***
				(-4.73)	(-4.82)	(-4.84)	(-4.70)	(-3.93)	(-3.92)
Job With New Employer				-0.043*	-0.043*	-0.047*	-0.049*	-0.041	-0.041
				(-2.00)	(-2.01)	(-2.17)	(-2.27)	(-1.86)	(-1.86)
Company Taken Over				0.158***	0.160***	0.159***	0.158***	0.168***	0.168***
				(5.61)	(5.70)	(5.68)	(5.63)	(5.95)	(5.94)
Changed Job, Same Firm				0.037	0.046	0.044	0.045	0.035	0.035
				(1.52)	(1.85)	(1.79)	(1.82)	(1.36)	(1.35)
FSIZE(Reference: GE 2000)									
[1] LT 20					0.038**	0.042**	0.045***	0.087***	0.087***
					(2.92)	(3.23)	(3.41)	(6.50)	(6.50)
[2] GE 20 LT 200					0.006	0.011	0.014	0.041**	0.041**
					(0.50)	(0.86)	(1.06)	(3.03)	(3.03)
[3] GE 200 LT 2000					0.002	0.005	0.006	0.017	0.017
					(0.14)	(0.37)	(0.43)	(1.14)	(1.13)
STATE(Reference: Bavaria)									
[1] Schleswig-Holstein						-0.028	-0.028	-0.011	-0.011
						(-1.02)	(-1.05)	(-0.42)	(-0.42)
[2] Hamburg						0.005	0.004	-0.002	-0.002
						(0.15)	(0.11)	(-0.05)	(-0.05)
[3] Lower Saxony						-0.021	-0.021	-0.009	-0.009
						(-1.23)	(-1.23)	(-0.52)	(-0.52)
[4] Bremen						-0.030	-0.029	-0.033	-0.033
						(-0.64)	(-0.60)	(-0.68)	(-0.68)
[5] North-Rhine-Westfalia						-0.012	-0.010	0.001	0.001
						(-0.81)	(-0.68)	(0.05)	(0.05)
[6] Hessen						0.020	0.019	0.020	0.020
						(1.05)	(1.04)	(1.07)	(1.07)
[7] Rheinland-Pfalz						-0.033	-0.035	-0.022	-0.022
						(-1.46)	(-1.55)	(-0.94)	(-0.94)
[8] Baden-Wuerttemberg						-0.005	-0.006	-0.004	-0.004
						(-0.33)	(-0.40)	(-0.22)	(-0.22)
[10] Saarland						-0.080	-0.079	-0.077	-0.077
						(-1.82)	(-1.79)	(-1.72)	(-1.72)

Continued on next page

Variables	Dependent variable: MATCH								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
[11] Berlin						-0.038	-0.040	-0.029	-0.029
						(-1.55)	(-1.64)	(-1.17)	(-1.17)
[12] Brandenburg						-0.092***	-0.090***	-0.065**	-0.065**
						(-4.05)	(-3.95)	(-2.82)	(-2.82)
[13]Mecklenburg-						-0.078**	-0.076**	-0.060*	-0.060*
Vorpommern						(-2.77)	(-2.72)	(-2.10)	(-2.10)
[14] Saxony						-0.007	-0.008	0.012	0.012
						(-0.39)	(-0.43)	(0.65)	(0.65)
[15] Saxony-						-0.049*	-0.047*	-0.024	-0.024
Anhalt						(-2.20)	(-2.12)	(-1.06)	(-1.06)
[16] Thuringia						-0.065**	-0.065**	-0.034	-0.034
						(-2.86)	(-2.83)	(-1.47)	(-1.48)
Survey year t (Reference: 2000)									
2001							0.017	0.015	0.015
							(0.88)	(0.76)	(0.76)
2002							0.038	0.037	0.037
							(1.86)	(1.77)	(1.77)
2003							0.021	0.016	0.016
							(0.96)	(0.73)	(0.73)
2004							0.021	0.023	0.023
							(0.98)	(1.06)	(1.06)
2005							0.035	0.032	0.032
							(1.62)	(1.44)	(1.43)
2006							0.007	0.007	0.007
							(0.34)	(0.33)	(0.33)
2007							0.018	0.022	0.022
							(0.86)	(1.06)	(1.06)
2008							0.029	0.033	0.033
							(1.38)	(1.55)	(1.55)
2009							0.058**	0.057**	0.057**
							(2.81)	(2.70)	(2.70)
2010							0.011	0.017	0.017
							(0.49)	(0.74)	(0.75)
2011							0.032	0.035	0.035
							(1.43)	(1.54)	(1.54)
2012							0.058*	0.069**	0.069**
							(2.51)	(2.94)	(2.95)
2013							0.061**	0.057*	0.057*
							(2.62)	(2.41)	(2.42)
2014							0.065**	0.077***	0.077***
							(3.23)	(3.77)	(3.78)
ISEI								0.0073***	0.0073***
								(21.08)	(21.08)
FEMALE									-0.0015
									(-0.16)
LR test(Prob> χ^2)		0.00	0.00	0.00	0.0026	0.0002	0.0521	0.00	0.8761
Observations	15779	15779	15779	15779	15779	15779	15779	15779	15779
Pseudo R^2	0.059	0.070	0.085	0.093	0.093	0.095	0.097	0.118	0.118

t statistics in parentheses. Marginal effects for factor levels is the discrete change from the base level.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix V. Estimation results: occupational mismatch using citizenship and search channels.

Table 3.17: Estimation results of occupational mismatch using citizenship and search channels.

Variables	Dependent variable: MATCH					
	(1)	(2)	(3)	(4)	(5)	(6)
EDU	0.027***	0.026***	0.026***	0.026***	0.025***	0.025***
	(12.50)	(12.18)	(11.95)	(11.95)	(11.54)	(11.51)
AGE	-0.006***	-0.006***	-0.006***	-0.006***	-0.006***	-0.006***
	(-13.95)	(-14.07)	(-14.18)	(14.18)	(-13.80)	(-13.82)
IND(Reference: Services)						
[1] Agriculture	0.032	0.034	0.033	0.033	0.031	0.032
	(0.97)	(1.04)	(1.01)	(1.01)	(0.95)	(0.99)
[2] Energy	-0.141**	-0.143**	-0.149***	-0.149***	-0.149***	-0.148***
	(-3.19)	(-3.24)	(-3.37)	(-3.37)	(-3.35)	(-3.34)
[3] Mining	-0.211	-0.207	-0.202	-0.202	-0.210	-0.209
	(-1.81)	(-1.77)	(-1.71)	(-1.71)	(-1.77)	(-1.76)
[4] Manufacturing	-0.043**	-0.042**	-0.040**	-0.040**	-0.040**	-0.039**
	(-3.28)	(-3.18)	(-3.05)	(-3.05)	(-3.03)	(-2.99)
[5] Construction	0.099***	0.101***	0.100***	0.100***	0.099***	0.099***
	(7.87)	(7.96)	(7.87)	(7.87)	(7.76)	(7.74)
[6] Trade	-0.109***	-0.108***	-0.104***	-0.104***	-0.106***	-0.106***
	(-8.92)	(-8.86)	(-8.49)	(-8.49)	(-8.63)	(-8.61)
[7] Transport	-0.202***	-0.199***	-0.198***	-0.198***	-0.200***	-0.200***
	(-10.84)	(-10.66)	(-10.57)	(-10.57)	(-10.66)	(-10.66)
[8] Bank,Insurance	-0.002	-0.003	-0.006	-0.006	-0.006	-0.006
	(-0.10)	(-0.13)	(-0.24)	(-0.24)	(-0.25)	(-0.25)
TOJCH(Reference: First job)						
Job After Break	-0.090***	-0.092***	-0.102***	-0.102***	-0.106***	-0.106***
	(-3.93)	(-4.00)	(-4.41)	(-4.41)	(-4.61)	(-4.60)
Job With New Employer	-0.041	-0.043	-0.041	-0.041	-0.040	-0.040
	(-1.86)	(-1.95)	(-1.84)	(-1.84)	(-1.83)	(-1.83)
Company Taken Over	0.168***	0.167***	0.147***	0.147***	0.134***	0.135***
	(5.95)	(5.89)	(5.08)	(5.08)	(4.56)	(4.57)
Changed Job, Same Firm	0.035	0.031	0.008	0.008	-0.015	-0.016
	(1.36)	(1.20)	(0.31)	(0.31)	(-0.59)	(-0.61)
FSIZE(Reference: GE 2000)						
[1] LT 20	0.087***	0.088***	0.098***	0.098***	0.102***	0.103***
	(6.50)	(6.53)	(7.28)	(7.28)	(7.56)	(7.60)
[2] GE 20 LT 200	0.041**	0.041**	0.045***	0.045***	0.050***	0.050***
	(3.03)	(3.03)	(3.30)	(3.30)	(3.63)	(3.66)
[3] GE 200 LT 2000	0.017	0.017	0.017	0.017	0.020	0.020
	(1.14)	(1.15)	(1.13)	(1.13)	(1.35)	(1.37)
STATE(Reference: Bavaria)						
[1] Schleswig-Holstein	-0.011	-0.018	-0.016	-0.016	-0.016	-0.016
	(-0.42)	(-0.66)	(-0.57)	(-0.57)	(-0.58)	(-0.59)
[2] Hamburg	-0.002	-0.006	-0.007	-0.007	-0.007	-0.007
	(-0.05)	(-0.17)	(-0.19)	(-0.19)	(-0.19)	(-0.19)
[3] Lower Saxony	-0.009	-0.013	-0.014	-0.014	-0.015	-0.015
	(-0.52)	(-0.75)	(-0.79)	(-0.79)	(-0.85)	(-0.84)
[4] Bremen	-0.033	-0.036	-0.037	-0.037	-0.034	-0.035
	(-0.68)	(-0.74)	(-0.75)	(-0.75)	(-0.70)	(-0.73)
[5] North-Rhine-Westfalia	0.001	0.000	0.001	0.001	0.000	0.000
	(0.05)	(0.01)	(0.08)	(0.08)	(0.02)	(0.01)
[6] Hessen	0.020	0.020	0.020	0.020	0.018	0.018
	(1.07)	(1.07)	(1.07)	(1.07)	(0.93)	(0.92)
[7] Rheinland-Pfalz	-0.021	-0.024	-0.024	-0.024	-0.025	-0.025
	(-0.94)	(-1.05)	(-1.02)	(-1.02)	(-1.07)	(-1.08)

Continued on next page

Variables	Dependent variable: MATCH					
	(1)	(2)	(3)	(4)	(5)	(6)
[8] Baden-Wuerttemberg	-0.004	0.001	-0.000	-0.000	-0.001	-0.000
	(-0.22)	(0.05)	(-0.02)	(-0.02)	(-0.03)	(-0.03)
[10] Saarland	-0.077	-0.075	-0.073	-0.073	-0.075	-0.073
	(-1.72)	(-1.67)	(-1.63)	(-1.63)	(-1.66)	(-1.63)
[11] Berlin	-0.029	-0.029	-0.028	-0.028	-0.029	-0.029
	(-1.17)	(-1.17)	(-1.15)	(-1.15)	(-1.16)	(-1.16)
[12] Brandenburg	-0.065**	-0.074**	-0.074**	-0.074**	-0.073**	-0.073**
	(-2.82)	(-3.18)	(-3.19)	(-3.19)	(-3.13)	(-3.14)
[13] Mecklenburg-Vorpommern	-0.060*	-0.069*	-0.072*	-0.072*	-0.069*	-0.070*
	(-2.10)	(-2.41)	(-2.50)	(-2.50)	(-2.42)	(-2.43)
[14] Saxony	0.012	0.004	0.003	0.003	0.003	0.002
	(0.65)	(0.21)	(0.14)	(0.14)	(0.15)	(0.13)
[15] Saxony-Anhalt	-0.024	-0.033	-0.034	-0.034	-0.031	-0.031
	(-1.06)	(-1.45)	(-1.49)	(-1.49)	(-1.36)	(-1.36)
[16] Thuringia	-0.034	-0.043	-0.040	-0.040	-0.039	-0.039
	(-1.47)	(-1.87)	(-1.71)	(-1.71)	(-1.66)	(-1.66)
Survey year t (Reference: 2000)						
2001	0.015	0.014	0.013	0.013	0.013	0.013
	(0.76)	(0.71)	(0.64)	(0.64)	(0.66)	(0.63)
2002	0.037	0.036	0.034	0.034	0.031	0.030
	(1.77)	(1.71)	(1.61)	(1.61)	(1.47)	(1.45)
2003	0.016	0.014	0.014	0.014	0.013	0.013
	(0.73)	(0.63)	(0.64)	(0.64)	(0.59)	(0.58)
2004	0.023	0.020	0.021	0.020	0.017	0.017
	(1.06)	(0.93)	(0.95)	(0.95)	(0.79)	(0.78)
2005	0.032	0.029	0.029	0.029	0.025	0.025
	(1.44)	(1.30)	(1.31)	(1.31)	(1.11)	(1.10)
2006	0.007	0.004	0.003	0.003	-0.001	-0.001
	(0.33)	(0.20)	(0.16)	(0.16)	(-0.03)	(-0.04)
2007	0.022	0.020	0.023	0.023	0.018	0.018
	(1.06)	(0.93)	(1.10)	(1.10)	(0.87)	(0.84)
2008	0.033	0.030	0.032	0.032	0.027	0.026
	(1.55)	(1.40)	(1.50)	(1.50)	(1.25)	(1.23)
2009	0.057**	0.055**	0.057**	0.057**	0.052*	0.052*
	(2.70)	(2.63)	(2.72)	(2.72)	(2.47)	(2.45)
2010	0.017	0.014	0.007	0.007	0.001	0.001
	(0.74)	(0.63)	(0.32)	(0.32)	(0.04)	(0.04)
2011	0.035	0.031	0.025	0.025	0.016	0.015
	(1.54)	(1.35)	(1.10)	(1.10)	(0.71)	(0.67)
2012	0.069**	0.065**	0.060*	0.060*	0.051*	0.052*
	(2.94)	(2.77)	(2.56)	(2.56)	(2.17)	(2.19)
2013	0.057*	0.053*	0.049*	0.049*	0.040	0.040
	(2.41)	(2.25)	(2.07)	(2.07)	(1.69)	(1.68)
2014	0.077***	0.077***	0.075***	0.075***	0.065**	0.064**
	(3.77)	(3.78)	(3.66)	(3.66)	(3.15)	(3.10)
ISEI	0.007***	0.007***	0.007***	0.007***	0.007***	0.007***
	(21.08)	(20.84)	(20.26)	(20.26)	(19.84)	(19.84)
Foreign citizen		-0.099***	-0.093***	-0.092***	-0.090***	0.003
		(-5.54)	(-5.14)	(-4.00)	(-5.00)	(0.04)
Referrals			-0.103***	-0.103***		
			(-10.87)	(-10.47)		
Foreign citizen × Referrals				-0.002		
				(-0.05)		
CHAN (Reference: Internet)						
Public emp. agency					-0.078***	-0.076***

Continued on next page

Variables	Dependent variable: MATCH					
	(1)	(2)	(3)	(4)	(5)	(6)
					(-3.70)	(-3.49)
Private emp. agency					-0.062	-0.060
					(-1.53)	(-1.44)
Newspaper					-0.056**	-0.056**
					(-2.83)	(-2.74)
Referrals					-0.129***	-0.125***
					(-7.33)	(-6.93)
Other					0.004	0.010
					(0.23)	(0.55)
MIG × CHAN(Reference: Foreign citizen × Internet)						
Foreign citizen × Public emp. agency						-0.073
						(-0.73)
Foreign citizen × Private emp. agency						-0.066
						(-0.41)
Foreign citizen × Newspaper						-0.048
						(-0.50)
Foreign citizen × Referrals						-0.098
						(-1.11)
Foreign citizen × Other						-0.137
						(-1.52)
LR test(Prob> χ^2)		0.00	0.00	0.9612	0.00	0.5276
Observations	15779	15779	15779	15779	15779	15779
Pseudo R^2	0.118	0.119	0.125	0.125	0.127	0.127

t statistics in parentheses. Marginal effects for factor levels is the discrete change from the base level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix VI. Estimation results: occupational mismatch using migration background and search channels.

Table 3.18: Estimation results of occupational mismatch using migration background and search channels.

Variables	Dependent variable: MATCH				
	(1)	(2)	(3)	(4)	(5)
EDU	0.027***	0.026***	0.026***	0.026***	0.025***
	(12.50)	(12.31)	(12.07)	(12.07)	(11.65)
AGE	-0.006***	-0.006***	-0.006***	-0.006***	-0.006***
	(-13.95)	(-13.83)	(-13.95)	(-13.95)	(-13.57)
IND(Reference: Services)					
[1] Agriculture	0.032	0.032	0.032	0.032	0.030
	(0.97)	(0.99)	(0.96)	(0.96)	(0.90)
[2] Energy	-0.141**	-0.142**	-0.147***	-0.147***	-0.148***
	(-3.19)	(-3.21)	(-3.34)	(-3.34)	(-3.33)
[3] Mining	-0.211	-0.211	-0.206	-0.206	-0.213
	(-1.81)	(-1.81)	(-1.75)	(-1.75)	(-1.81)
[4] Manufacturing	-0.043**	-0.041**	-0.039**	-0.039**	-0.039**
	(-3.28)	(-3.09)	(-2.97)	(-2.98)	(-2.96)
[5] Construction	0.099***	0.101***	0.100***	0.100***	0.099***
	(7.87)	(7.98)	(7.89)	(7.89)	(7.78)
[6] Trade	-0.109***	-0.108***	-0.104***	-0.104***	-0.106***
	(-8.92)	(-8.83)	(-8.48)	(-8.48)	(-8.62)

Continued on next page

Variables	Dependent variable: MATCH				
	(1)	(2)	(3)	(4)	(5)
[7] Transport	-0.202***	-0.199***	-0.198***	-0.198***	-0.200***
	(-10.84)	(-10.64)	(-10.56)	(-10.56)	(-10.65)
[8] Bank,Insurance	-0.002	-0.005	-0.008	-0.008	-0.008
	(-0.10)	(-0.21)	(-0.31)	(-0.31)	(-0.31)
TOJCH(Reference: First job)					
Job After Break	-0.090***	-0.092***	-0.102***	-0.102***	-0.106***
	(-3.93)	(-4.01)	(-4.42)	(-4.42)	(-4.61)
Job With New Employer	-0.041	-0.044	-0.042	-0.042	-0.041
	(-1.86)	(-1.99)	(-1.88)	(-1.88)	(-1.86)
Company Taken Over	0.168***	0.166***	0.147***	0.147***	0.134***
	(5.95)	(5.86)	(5.06)	(5.06)	(4.55)
Changed Job, Same Firm	0.035	0.030	0.007	0.007	-0.015
	(1.36)	(1.17)	(0.29)	(0.29)	(-0.58)
FSIZE(Reference: GE 2000)					
[1] LT 20	0.087***	0.087***	0.098***	0.098***	0.102***
	(6.50)	(6.51)	(7.26)	(7.26)	(7.54)
[2] GE 20 LT 200	0.041**	0.042**	0.046***	0.046***	0.050***
	(3.03)	(3.08)	(3.35)	(3.35)	(3.68)
[3] GE 200 LT 2000	0.017	0.017	0.017	0.017	0.020
	(1.14)	(1.15)	(1.14)	(1.14)	(1.35)
STATE(Reference: Bavaria)					
[1] Schleswig-Holstein	-0.011	-0.016	-0.014	-0.014	-0.014
	(-0.42)	(-0.60)	(-0.51)	(-0.51)	(-0.52)
[2] Hamburg	-0.002	-0.004	-0.005	-0.005	-0.005
	(-0.05)	(-0.12)	(-0.14)	(-0.13)	(-0.15)
[3] Lower Saxony	-0.009	-0.010	-0.011	-0.011	-0.012
	(-0.52)	(-0.58)	(-0.63)	(-0.63)	(-0.69)
[4] Bremen	-0.033	-0.028	-0.029	-0.029	-0.027
	(-0.68)	(-0.57)	(-0.60)	(-0.60)	(-0.55)
[5] North-Rhine-Westfalia	0.001	0.003	0.003	0.003	0.002
	(0.05)	(0.18)	(0.23)	(0.23)	(0.16)
[6] Hessen	0.020	0.022	0.022	0.022	0.020
	(1.07)	(1.19)	(1.17)	(1.17)	(1.03)
[7] Rheinland-Pfalz	-0.021	-0.020	-0.020	-0.020	-0.021
	(-0.94)	(-0.87)	(-0.85)	(-0.85)	(-0.91)
[8] Baden-Wuerttemberg	-0.004	0.002	0.001	0.001	0.000
	(-0.22)	(0.13)	(0.05)	(0.05)	(0.03)
[10] Saarland	-0.077	-0.069	-0.067	-0.067	-0.069
	(-1.72)	(-1.53)	(-1.49)	(-1.49)	(-1.53)
[11] Berlin	-0.029	-0.028	-0.028	-0.028	-0.028
	(-1.17)	(-1.14)	(-1.11)	(-1.12)	(-1.14)
[12] Brandenburg	-0.065**	-0.075**	-0.075**	-0.075**	-0.074**
	(-2.82)	(-3.24)	(-3.24)	(-3.24)	(-3.18)
[13] Mecklenburg-Vorpommern	-0.060*	-0.071*	-0.074**	-0.074**	-0.071*
	(-2.10)	(-2.49)	(-2.57)	(-2.58)	(-2.48)
[14] Saxony	0.012	0.002	0.001	0.001	0.001
	(0.65)	(0.10)	(0.04)	(0.04)	(0.06)
[15] Saxony-Anhalt	-0.024	-0.035	-0.036	-0.036	-0.033
	(-1.06)	(-1.53)	(-1.56)	(-1.56)	(-1.42)
[16] Thuringia	-0.034	-0.047*	-0.043	-0.043	-0.041
	(-1.47)	(-2.00)	(-1.83)	(-1.83)	(-1.77)
Survey year t (Reference: 2000)					
2001	0.015	0.015	0.013	0.013	0.014
	(0.76)	(0.75)	(0.67)	(0.67)	(0.69)
2002	0.037	0.036	0.034	0.034	0.031

Continued on next page

Variables	Dependent variable: MATCH				
	(1)	(2)	(3)	(4)	(5)
	(1.77)	(1.71)	(1.60)	(1.60)	(1.46)
2003	0.016	0.015	0.015	0.015	0.014
	(0.73)	(0.66)	(0.66)	(0.66)	(0.62)
2004	0.023	0.023	0.023	0.023	0.019
	(1.06)	(1.04)	(1.05)	(1.05)	(0.89)
2005	0.032	0.030	0.031	0.030	0.026
	(1.44)	(1.37)	(1.37)	(1.37)	(1.17)
2006	0.007	0.006	0.005	0.005	0.001
	(0.33)	(0.30)	(0.24)	(0.24)	(0.05)
2007	0.022	0.022	0.025	0.025	0.020
	(1.06)	(1.05)	(1.21)	(1.21)	(0.96)
2008	0.033	0.033	0.035	0.034	0.029
	(1.55)	(1.54)	(1.63)	(1.63)	(1.37)
2009	0.057**	0.058**	0.060**	0.060**	0.054*
	(2.70)	(2.76)	(2.85)	(2.85)	(2.58)
2010	0.017	0.017	0.010	0.010	0.003
	(0.74)	(0.75)	(0.43)	(0.43)	(0.14)
2011	0.035	0.035	0.029	0.029	0.020
	(1.54)	(1.53)	(1.26)	(1.26)	(0.86)
2012	0.069**	0.067**	0.063**	0.063**	0.054*
	(2.94)	(2.88)	(2.67)	(2.66)	(2.27)
2013	0.057*	0.056*	0.052*	0.052*	0.043
	(2.41)	(2.37)	(2.18)	(2.18)	(1.79)
2014	0.077***	0.082***	0.079***	0.079***	0.069***
	(3.77)	(4.04)	(3.89)	(3.89)	(3.37)
ISEI	0.007***	0.007***	0.007***	0.007***	0.007***
	(21.08)	(20.61)	(20.06)	(20.05)	(19.65)
MIGBACK (Reference: German national)					
Direct migrant		-0.087***	-0.080***	-0.079***	-0.077***
		(-5.81)	(-5.33)	(-4.19)	(-5.15)
Indirect migrant		-0.029	-0.027	-0.030	-0.026
		(-1.65)	(-1.50)	(-1.37)	(-1.44)
Referrals			-0.103***	-0.103***	
			(-10.81)	(-9.89)	
MIGBACK × REF (Reference: German national × Formal channels)					
Direct migrant × referrals				-0.002	
				(-0.06)	
Indirect migrant × referrals				0.009	
				(0.25)	
Chan (Reference: Internet)					
Public emp. agency					-0.079***
					(-3.75)
Private emp. agency					-0.061
					(-1.51)
Newspaper					-0.058**
					(-2.92)
Referrals					-0.130***
					(-7.39)
Other					0.002
					(0.12)
LR test(Prob> χ^2)		0.00	0.00	0.9666	0.00
Observations	15779	15779	15779	15779	15779
Pseudo R^2	0.118	0.119	0.125	0.125	0.127

t statistics in parentheses. Marginal effects for factor levels is the discrete change from the base level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix VII. Robustness check: empirical approach.

Since the dependent variable is binary, we use binary choice regression models in empirical estimations. Moreover, since the data used is longitudinal, we consider logit and probit fixed and random effects regression models to deal with issues which may be caused by the potential endogenous variables in the regressions.

Estimation of fixed effects logit and probit models will raise an incidental parameter problem in case of the number of time periods T is fixed. Thus, introducing inconsistency in the estimators of the constants terms, leading to an inconsistent coefficient of maximum-likelihood estimator (MLE) since MLE is a function of the estimators of the constant terms. (see Neyman et al. (1948) and Lancaster (2000).) Yet, following Rasch (1960) and Andersen (1970), a concept of conditional maximum likelihood estimator (CMLE) is proposed by Chamberlain (1980). The CMLE is consistent given the conditional likelihood function satisfies regularity conditions. According to the regularity conditions, mild restrictions are imposed on the incidental parameters discussed in Andersen (1970), Andersen (1971). Chamberlain (1980) demonstrates that likelihood function is free of the incidental parameters conditional on sufficient statistics for the incidental parameters. Therefore, the CMLE is computationally convenient estimator for the fixed effects logit model because in this model the sum of the individual dependent variable's value over time is a minimal sufficient statistic for the individual constant term. However, this approach is not free of incidental parameters problem for the fixed effects probit model, because the incidental parameters of the fixed effects probit model cannot be removed from the conditional likelihood function, because there are no sufficient statistics available for the corresponding probit model. Nonetheless, our estimations do not favor conditional ML estimation of fixed effects logit model for the following reasons. First, since our individual dependent variables often do not change over time the conditional probabilities of these observations contribute nothing to the conditional likelihood function. As a result, CMLE estimations use only 3308 observations out of 15779. Second, fixed effects are not estimated, which in turn prevents the estimation of marginal effects with the coefficients estimated by CMLE.

Though fixed effects logit model is favored over probit model, the case is different for the random effects models. This is driven by the important distinction between the assumptions of error distributions in the two models mentioned. Logit model uses multivariate logistic have a logistic distribution, because in the logit model errors are assumed to have a logistic distribution. The disadvantage of the multivariate logistic distribution is that the correlations are all constrained to be 0.5. The probit model is more flexible, because it is based on the multivariate normal distribution, more flexible. (see Johnson and Kotz (1972), Maddala (1987).) The drawback of binary choice random effects models compared to the binary choice fixed 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.) The model can be extended to binary choice setting using the method specified by Butler and Moffitt (1982). Afterwords, an approximation of log likelihood can be obtained using a Gauss-Hermite quadrature technique. Estimation of the random effects probit model was conducted using the statistical program Stata, where the panel-level likelihood is approximated following adaptive Gauss-Hermite quadrature method of Naylor and Smith (1982).

Appendix VIII. Individual effects estimation results: occupational mismatch.

Using panel probit regression with random effects, $MATCH_{i,t}$ is regressed sequentially on different control variables. As before we conduct the likelihood-ratio test for each set of control variables. The corresponding regression output and likelihood ratios are presented in Table 3.19. The results of likelihood-ratio tests suggest that among the control variables only the dummy variable indicating gender of the individual should not be added to the regression equation. Our results reveal that higher education is positively associated with the probability of a good match. At the same time we can see that workers in smaller firms are more likely to perform a job corresponding to their initial training, whereas workers in larger firms are more frequently mismatched. Furthermore, jobs obtained after a long break are often associated with mismatch. These results are qualitatively similar to the results presented in Appendix IV, where logistic regression model is used in estimations.

Moreover, we compare the panel probit estimator to the pooled probit estimator using ρ, which is the proportion of the total variance contributed by the panel-level variance component. A likelihood-ratio test is conducted to check if the panel estimator is different from the pooled estimator. The tests for all the specifications (1) to (9) suggest that the panel-level variance component is significantly more than zero, which implies that panel-level variance is important and panel estimator is different from the pooled estimator.

Table 3.19: Estimation results of occupational mismatch.

Variables	Dependent variable: MATCH								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EDU	0.096***	0.095***	0.090***	0.086***	0.087***	0.087***	0.087***	0.050***	0.050***
	(26.59)	(26.59)	(25.05)	(23.94)	(23.99)	(23.81)	(23.61)	(12.90)	(12.90)
AGE		-0.0098***	-0.0098***	-0.0088***	-0.0089***	-0.0088***	-0.0095***	-0.0092***	-0.0092***
		(-13.57)	(-13.67)	(-12.10)	(-12.19)	(-12.03)	(-12.46)	(-12.31)	(-12.31)
IND(Reference: Services)									
[1] Agriculture			-0.029	-0.021	-0.032	-0.018	-0.019	0.060	0.060
			(-0.55)	(-0.40)	(-0.61)	(-0.34)	(-0.36)	(1.25)	(1.24)
[2] Energy			-0.148*	-0.160*	-0.150*	-0.149*	-0.156*	-0.185**	-0.185**
			(-2.20)	(-2.36)	(-2.20)	(-2.19)	(-2.28)	(-2.74)	(-2.74)
[3] Mining			-0.344 *	-0.364*	-0.350*	-0.331*	-0.338*	-0.329*	-0.329*
			(-2.15)	(-2.37)	(-2.23)	(-2.05)	(-2.10)	(-2.05)	(-2.05)
[4] Manufacturing			-0.073***	-0.074***	-0.068**	-0.073***	-0.068**	-0.059**	-0.059**
			(-4.01)	(-3.57)	(-3.26)	(-3.50)	(-3.24)	(-2.81)	(-2.74)

Continued on next page

Variables	Dependent variable: MATCH								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
[5] Construction			0.122***	0.124***	0.124***	0.126***	0.131***	0.146***	0.146***
			(6.55)	(6.70)	(6.68)	(6.79)	(7.02)	(8.11)	(7.77)
[6] Trade			-0.130***	-0.122***	-0.128***	-0.130***	-0.128***	-0.137***	-0.137***
			(-6.80)	(-6.36)	(-6.63)	(-6.76)	(-6.57)	(-7.07)	(-7.07)
[7] Transport			-0.284***	-0.289***	-0.288***	-0.287***	-0.284***	-0.264***	-0.264***
			(-9.68)	(-9.87)	(-9.76)	(-9.72)	(-9.56)	(-8.86)	(-8.76)
[8] Bank,Insurance			0.083*	0.070*	0.077*	0.070*	0.078*	0.020	0.020
			(2.43)	(1.99)	(2.17)	(1.97)	(2.19)	(0.54)	(0.54)
TOJCH(Reference: First job)									
Job After Break				-0.139***	-0.142***	-0.141***	-0.137***	-0.121***	-0.121***
				(-4.36)	(-4.42)	(-4.41)	(-4.27)	(-3.72)	(-3.72)
Job With New Employer				-0.074*	-0.074*	-0.077*	-0.086**	-0.077*	-0.077*
				(-2.43)	(-2.42)	(-2.53)	(-2.81)	(-2.47)	(-2.47)
Company Taken Over				0.196***	0.197***	0.197***	0.192***	0.200***	0.200***
				(5.65)	(5.69)	(5.71)	(5.56)	(5.74)	(5.74)
Changed Job, Same Firm				0.039	0.046	0.045	0.043	0.030	0.030
				(1.14)	(1.31)	(1.27)	(1.22)	(0.84)	(0.84)
FSIZE(Reference: GE 2000)									
[1] LT 20					0.040*	0.044*	0.050*	0.104***	0.104***
					(2.01)	(2.20)	(2.48)	(5.14)	(5.13)
[2] GE 20 LT 200					0.010	-0.006	-0.000	0.037	0.037
					(0.51)	(-0.28)	(-0.01)	(1.80)	(1.80)
[3] GE 200 LT 2000					-0.027	-0.024	-0.023	-0.006	-0.006
					(-1.23)	(-1.08)	(-1.04)	(-0.29)	(-0.29)
STATE(Reference: Bavaria)									
[1] Schleswig-Holstein						-0.068	-0.069	-0.043	-0.043
						(-1.39)	(-1.39)	(-0.88)	(-0.88)
[2] Hamburg						0.007	0.007	-0.007	-0.007
						(0.12)	(0.11)	(-0.12)	(-0.12)
[3] Lower Saxony						-0.026	-0.024	-0.004	-0.004
						(-0.83)	(-0.77)	(-0.14)	(-0.14)
[4] Bremen						-0.075	-0.071	-0.067	-0.067
						(-0.85)	(-0.80)	(-0.77)	(-0.77)
[5] North-Rhine-Westfalia						-0.014	-0.010	0.005	0.005
						(-0.56)	(-0.38)	(0.19)	(0.19)
[6] Hessen						-0.003	-0.003	0.001	0.001
						(-0.09)	(-0.09)	(0.04)	(0.04)
[7] Rheinland-Pfalz						-0.038	-0.040	-0.019	-0.019
						(-0.92)	(-0.96)	(-0.46)	(-0.46)
[8] Baden-Wuerttemberg						0.015	0.013	0.016	0.016
						(0.52)	(0.44)	(0.56)	(0.56)
[10] Saarland						-0.118	-0.113	-0.108	-0.108
						(-1.46)	(-1.39)	(-1.34)	(-1.34)
[11] Berlin						-0.076	-0.080	-0.063	-0.063
						(-1.71)	(-1.78)	(-1.42)	(-1.42)
[12] Brandenburg						-0.144***	-0.137**	-0.096*	-0.096**
						(-3.42)	(-3.24)	(-2.29)	(-2.29)
[13]Mecklenburg-Vorpommern						-0.124*	-0.121*	-0.094	-0.094
						(-2.33)	(-2.24)	(-1.78)	(-1.77)
[14] Saxony						-0.016	-0.016	0.013	0.013
						(-0.48)	(-0.46)	(0.40)	(0.40)
[15] Saxony-Anhalt						-0.103*	-0.098*	-0.058	-0.058
						(-2.41)	(-2.28)	(-1.38)	(-1.38)
[16] Thuringia						-0.114**	-0.113**	-0.065	-0.065
						(-2.72)	(-2.66)	(-1.56)	(-1.56)

Continued on next page

Variables	Dependent variable: MATCH								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Survey year t (Reference: 2000)									
2001							0.035	0.028	0.028
							(1.22)	(0.97)	(0.97)
2002							0.045	0.041	0.041
							(1.49)	(1.35)	(1.35)
2003							0.026	0.018	0.018
							(0.81)	(0.58)	(0.58)
2004							0.025	0.025	0.025
							(0.77)	(0.80)	(0.80)
2005							0.059	0.053	0.053
							(1.85)	(1.67)	(1.67)
2006							0.012	0.015	0.015
							(0.38)	(0.47)	(0.047)
2007							0.010	0.013	0.013
							(0.34)	(0.44)	(0.44)
2008							0.056	0.056	0.056
							(1.83)	(1.84)	(1.84)
2009							0.090**	0.082**	0.082**
							(2.97)	(2.72)	(2.72)
2010							0.039	0.036	0.036
							(1.18)	(1.11)	(1.11)
2011							0.076*	0.072*	0.072*
							(2.32)	(2.20)	(2.20)
2012							0.097**	0.104**	0.104**
							(2.89)	(3.13)	(3.13)
2013							0.115***	0.105**	0.105**
							(3.44)	(3.15)	(3.14)
2014							0.118**	0.126***	0.126***
							(3.92)	(4.24)	(4.23)
ISEI								0.0107***	0.0107***
								(18.99)	(18.99)
FEMALE									0.0003
									(0.02)
ρ	0.680	0.668	0.658	0.653	0.655	0.654	0.658	0.642	0.642
Std. Err. of ρ	(0.014)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
χ^2	1540.99	1476.79	1402.85	1374.04	1375.75	1368.03	1378.11	1307.16	1307.13
Prob> χ^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LR test(Prob> χ^2)		0.00	0.00	0.00	0.0015	0.0026	0.0026	0.00	0.9856
Observations	15779	15779	15779	15779	15779	15779	15779	15779	15779
Pseudo R^2	0.130	0.139	0.149	0.156	0.157	0.159	0.160	0.178	0.178

t statistics in parentheses. Marginal effects for factor levels is the discrete change from the base level.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix IX. Individual effects estimation results: occupational mismatch using citizenship and search channels.

Furthermore, $MIG_{i,t}$ is added to the regression equation of panel probit model with random effects. The marginal effects are presented in Table 3.20 and the t statistics are contained in brackets. The results of panel probit regression with random effects confirm all the conclusions derived from the results of logistic regression with the same set of independent variables presented in Table 3.10. Namely, foreign workers have 15.4% lower probability of being well matched in the job, see column (2). Using the same set of independent variable logistic regression results from Table 3.10 show that foreign workers have 10% lower probability of being well matched in the job. According to column (3) the fact that foreign workers rely more often on referral hiring explains a part of the negative link (0.8%) between being a foreigner and the probability of a good match. The corresponding part of the negative link is 0.7% in case of using logistic regression. The likelihood-ratio test suggests that $MIG_{i,t} \times REF_{i,t}$ should not be included into the regression equation since this variable is not significant, see column (4). This confirms that referrals have equally low efficiency in generating good matches irrespective of the applicant's ethnic belonging.

In the next step, $CHAN_{i,t}$ is added to the regression equation instead of $REF_{i,t}$. Again, referrals lead most often to mismatch compared to other search channels. Other search channels which are positively associated with mismatch are newspapers and the public employment agency. When detailed information about the search channel is used we can see that the marginal effect of variable MIG is reduced even further from 14.6% down to 14.2%.The fact that foreign workers rely more often on newspapers and the public employment agency explains another 0.4% difference in the probability of mismatch between native and foreign workers (see column (5)). The corresponding additionally explained part is 0.3% in case of using logistic regression. Finally, in specification (6) we additionally include the interaction terms between $MIG_{i,t}$ and $CHAN_{i,t}$, but none of these interaction terms is statistically significant. Moreover, the likelihood-ratio test suggests that the interaction terms should not be included to the regression equation. Again this shows that different search channels have similar match qualities when used by native and foreign workers. It is rather so that foreign workers are more likely to rely on search channels with lower efficiency, like referral hiring and employment agency.

We again compare the panel probit estimator to the pooled probit estimator for each specification. The tests suggest that the panel-level variance component is significantly more than zero, which implies that panel-level variance is important in estimations of all the regression equations.

Table 3.20: Estimation results of occupational mismatch using citizenship and search channels.

Variables	Dependent variable: MATCH					
	(1)	(2)	(3)	(4)	(5)	(6)
EDU	0.050***	0.048***	0.047***	0.047***	0.046***	0.046***
	(12.90)	(12.59)	(12.38)	(12.38)	(12.03)	(12.01)
AGE	-0.0092***	-0.0093***	-0.0093***	-0.0093***	-0.0091***	-0.0091***
	(-12.31)	(-12.42)	(-12.49)	(-12.48)	(-12.23)	(-12.23)
IND(Reference: Services)						
[1] Agriculture	0.060	0.062	0.065	0.065	0.060	0.061
	(1.25)	(1.31)	(1.37)	(1.37)	(1.24)	(1.27)
[2] Energy	-0.185**	-0.187**	-0.195**	-0.196**	-0.199**	-0.198**
	(-2.74)	(-2.78)	(-2.91)	(-2.91)	(-2.96)	(-2.95)
[3] Mining	-0.329*	-0.324*	-0.296	-0.297	-0.309	-0.310
	(-2.05)	(-2.01)	(-1.81)	(-1.81)	(-1.91)	(-1.92)

Continued on next page

Variables	Dependent variable: MATCH					
	(1)	(2)	(3)	(4)	(5)	(6)
[4] Manufacturing	-0.059**	-0.057**	-0.054**	-0.054**	-0.054**	-0.054**
	(-2.81)	(-2.72)	(-2.59)	(-2.58)	(-2.60)	(-2.58)
[5] Construction	0.146***	0.147***	0.147***	0.147***	0.145***	0.144***
	(8.11)	(8.18)	(8.14)	(8.14)	(8.01)	(7.98)
[6] Trade	-0.137***	-0.136***	-0.132***	-0.132***	-0.135***	-0.135***
	(-7.07)	(-7.03)	(-6.84)	(-6.83)	(-6.98)	(-6.95)
[7] Transport	-0.264***	-0.260***	-0.257***	-0.257***	-0.261***	-0.261***
	(-8.86)	(-8.72)	(-8.59)	(-8.58)	(-8.72)	(-8.72)
[8] Bank,Insurance	0.020	0.019	0.015	0.015	0.014	0.014
	(0.54)	(0.49)	(0.40)	(0.40)	(0.38)	(0.37)
TOJCH(Reference: First job)						
Job After Break	-0.121***	-0.123***	-0.134***	-0.134***	-0.138***	-0.137***
	(-3.72)	(-3.79)	(-4.13)	(-4.13)	(-4.27)	(-4.26)
Job With New Employer	-0.077*	-0.079*	-0.074*	-0.073*	-0.073*	-0.073*
	(-2.47)	(-2.56)	(-2.39)	(-2.39)	(-2.39)	(-2.40)
Company Taken Over	0.200***	0.197***	0.179***	0.179***	0.167***	0.167***
	(5.74)	(5.68)	(5.02)	(5.02)	(4.62)	(4.62)
Changed Job, Same Firm	0.030	0.025	-0.002	-0.002	-0.027	-0.029
	(0.84)	(0.70)	(-0.06)	(-0.07)	(-0.73)	(-0.77)
FSIZE(Reference: GE 2000)						
[1] LT 20	0.104***	0.104***	0.117***	0.117***	0.122***	0.122***
	(5.14)	(5.16)	(5.78)	(5.78)	(6.00)	(6.01)
[2] GE 20 LT 200	0.037	0.037	0.043*	0.043*	0.048*	0.048*
	(1.80)	(1.81)	(2.10)	(2.10)	(2.36)	(2.36)
[3] GE 200 LT 2000	-0.006	-0.006	-0.007	-0.007	-0.002	-0.002
	(-0.29)	(-0.29)	(-0.30)	(-0.30)	(-0.09)	(-0.08)
STATE(Reference: Bavaria)						
[1] Schleswig-Holstein	-0.043	-0.053	-0.048	-0.048	-0.049	-0.049
	(-0.88)	(-1.07)	(-0.99)	(-0.98)	(-1.00)	(-1.01)
[2] Hamburg	-0.007	-0.014	-0.018	-0.018	-0.020	-0.020
	(-0.12)	(-0.22)	(-0.30)	(-0.30)	(-0.33)	(-0.32)
[3] Lower Saxony	-0.004	-0.010	-0.010	-0.010	-0.011	-0.011
	(-0.14)	(-0.32)	(-0.32)	(-0.33)	(-0.36)	(-0.36)
[4] Bremen	-0.067	-0.069	-0.066	-0.066	-0.064	-0.066
	(-0.77)	(-0.79)	(-0.77)	(-0.76)	(-0.74)	(-0.76)
[5] North-Rhine-Westfalia	0.005	0.004	0.005	0.005	0.004	0.004
	(0.19)	(0.16)	(0.20)	(0.20)	(0.16)	(0.16)
[6] Hessen	0.001	0.002	0.002	0.001	-0.001	-0.002
	(0.04)	(0.06)	(0.05)	(0.04)	(-0.04)	(-0.06)
[7] Rheinland-Pfalz	-0.019	-0.022	-0.022	-0.022	-0.023	-0.023
	(-0.46)	(-0.55)	(-0.54)	(-0.54)	(-0.57)	(-0.58)
[8] Baden-Wuerttemberg	0.016	0.023	0.020	0.021	0.020	0.020
	(0.56)	(0.80)	(0.71)	(0.72)	(0.69)	(0.68)
[10] Saarland	-0.108	-0.102	-0.099	-0.099	-0.102	-0.098
	(-1.34)	(-1.27)	(-1.25)	(-1.24)	(-1.28)	(-1.24)
[11] Berlin	-0.063	-0.063	-0.061	-0.061	-0.062	-0.062
	(-1.42)	(-1.43)	(-1.38)	(-1.38)	(-1.40)	(-1.41)
[12] Brandenburg	-0.096*	-0.109**	-0.109**	-0.109**	-0.107*	-0.108**
	(-2.29)	(-2.60)	(-2.62)	(-2.62)	(-2.57)	(-2.58)
[13] Mecklenburg-	-0.094	-0.107*	-0.110*	-0.110*	-0.106*	-0.107*
Vorpommern	(-1.78)	(-2.03)	(-2.09)	(-2.09)	(-2.02)	(-2.03)
[14] Saxony	0.013	0.001	-0.001	-0.001	0.001	-0.000
	(0.40)	(0.04)	(-0.02)	(-0.02)	(0.02)	(-0.00)
[15] Saxony-Anhalt	-0.058	-0.072	-0.071	-0.071	-0.067	-0.067
	(-1.38)	(-1.70)	(-1.68)	(-1.68)	(-1.58)	(-1.60)

Continued on next page

Variables	Dependent variable: MATCH					
	(1)	(2)	(3)	(4)	(5)	(6)
[16] Thuringia	-0.065	-0.079	-0.073	-0.073	-0.071	-0.071
	(-1.56)	(-1.88)	(-1.76)	(-1.76)	(-1.71)	(-1.71)
Survey year t (Reference: 2000)						
2001	0.028	0.027	0.026	0.026	0.027	0.026
	(0.97)	(0.93)	(0.92)	(0.93)	(0.94)	(0.92)
2002	0.041	0.040	0.038	0.038	0.034	0.034
	(1.35)	(1.31)	(1.24)	(1.24)	(1.13)	(1.11)
2003	0.018	0.016	0.017	0.017	0.015	0.015
	(0.58)	(0.50)	(0.55)	(0.55)	(0.48)	(0.47)
2004	0.025	0.022	0.023	0.023	0.018	0.018
	(0.80)	(0.69)	(0.72)	(0.72)	(0.58)	(0.57)
2005	0.053	0.049	0.051	0.051	0.046	0.046
	(1.67)	(1.55)	(1.60)	(1.60)	(1.44)	(1.44)
2006	0.015	0.011	0.012	0.012	0.007	0.007
	(0.47)	(0.36)	(0.39)	(0.39)	(0.23)	(0.23)
2007	0.013	0.010	0.017	0.017	0.010	0.010
	(0.44)	(0.33)	(0.54)	(0.55)	(0.32)	(0.31)
2008	0.056	0.052	0.057	0.057	0.050	0.050
	(1.84)	(1.72)	(1.87)	(1.88)	(1.66)	(1.65)
2009	0.082**	0.080**	0.083**	0.083**	0.076*	0.076*
	(2.72)	(2.65)	(2.75)	(2.75)	(2.52)	(2.51)
2010	0.036	0.033	0.023	0.024	0.013	0.013
	(1.11)	(1.01)	(0.72)	(0.72)	(0.41)	(0.40)
2011	0.072*	0.067*	0.060	0.060	0.048	0.047
	(2.20)	(2.04)	(1.83)	(1.83)	(1.44)	(1.42)
2012	0.104**	0.099**	0.095**	0.095**	0.084*	0.084*
	(3.13)	(2.98)	(2.84)	(2.84)	(2.48)	(2.50)
2013	0.105**	0.101**	0.096**	0.096**	0.085*	0.085*
	(3.15)	(3.01)	(2.86)	(2.86)	(2.52)	(2.51)
2014	0.126***	0.126***	0.123***	0.123***	0.110***	0.109***
	(4.24)	(4.25)	(4.14)	(4.14)	(3.67)	(3.64)
ISEI	0.0107***	0.0106***	0.0103***	0.0103***	0.0101***	0.0101***
	(18.99)	(18.81)	(18.38)	(18.38)	(18.06)	(18.06)
Foreign		-0.154***	-0.146***	-0.155***	-0.142***	0.005
citizen		(-4.76)	(-4.52)	(-3.99)	(-4.41)	(0.04)
Referrals			-0.138***	-0.140***		
			(-9.77)	(-9.53)		
Foreign citizen ×				0.021		
Referrals				(0.41)		
CHAN (Reference: Internet)						
Public emp. agency					-0.106***	-0.101**
					(-3.55)	(-3.28)
Private emp. agency					-0.076	-0.081
					(-1.35)	(-1.38)
Newspaper					-0.073**	-0.072*
					(-2.65)	(-2.54)
Referrals					-0.174***	-0.170***
					(-7.09)	(-6.74)
Other					-0.001	0.007
					(-0.05)	(0.28)
MIG × CHAN(Reference: Foreign citizen × Internet)						
Foreign citizen × Public						-0.150
emp. agency						(-0.99)
Foreign citizen × Private						-0.015
emp. agency						(-0.06)

Continued on next page

Variables	Dependent variable: MATCH					
	(1)	(2)	(3)	(4)	(5)	(6)
Foreign citizen ×						-0.095
Newspaper						(-0.65)
Foreign citizen × Referrals						-0.141
						(-1.03)
Foreign citizen × Other						-0.217
						(-1.59)
ρ	0.642	0.641	0.638	0.638	0.637	0.637
Std. Err. of ρ	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
χ^2	1307.16	1300.23	1278.76	1278.91	1270.06	1270.32
Prob> χ^2	0.00	0.00	0.00	0.00	0.00	0.00
LR test(Prob> χ^2)		0.00	0.00	0.6861	0.00	0.5037
Observations	15779	15779	15779	15779	15779	15779
Pseudo R^2	0.178	0.180	0.184	0.184	0.185	0.186

t statistics in parentheses. Marginal effects for factor levels is the discrete change from the base level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix X. Individual effects estimation results: occupational mismatch using migration background and search channels.

In Table 3.21 we substitute binary variable $MIG_{i,t}$ with a more detailed variable $MIGBACK_{i,t}$ containing three categories. The results of panel probit regression with random effects again confirm all the conclusions derived from the results of logistic regression with the same set of independent variables presented in Table 3.11. Column (2) shows that compared to German nationals direct migrants are less likely to be well matched, while indirect migrants can not be statistically distinguished from native German workers. The marginal effect shows that direct migrants are 14.2% more likely to be mismatched than German nationals. Next, $REF_{i,t}$ is added to the regression in column (3). The negative and statistically significant coefficient of referrals suggests that referral hiring leads to good matches less often compared to hiring through formal search channels. We can see that the marginal effect is again reduced from 14.2% down to 13.2%. This confirms our earlier conclusion that 1% of the differences in mismatch rates between migrant and native workers is due to the fact that migrants rely more often on their social networks. Now we can additionally conclude that this effect is largely generated by direct migrants. The corresponding part of the explained difference in case of using logistic regression is 0.7%. The interaction terms in column (4) are again insignificant.

Next, a more detailed variable $CHAN_{i,t}$ for the search channel is included instead of a binary indicator $REF_{i,t}$. The marginal effect of being a direct migrant falls from 13.2% down to 12.9%, so this regression confirms the fact that additional 0.3% difference in the probabilities of mismatch is due to the fact that direct migrants use less efficient search channels such as newspapers and services of the public employment agency more often than native German workers. According to logistic regression results again additional 0.3% difference in the probabilities of mismatch is due to the fact that direct migrants use less efficient search channels. The results with interaction terms between the search channels and $MIGBACK_{i,t}$ are not included in Table 3.21 as none of the interaction terms was significant in the previous regressions.

Similar to the previous panel probit estimations, for all the five specifications the likelihood-ratio tests suggest that the panel-level variance component is significantly more than zero.

Table 3.21: Estimation results of occupational mismatch using migration background and search channels.

Variables	Dependent variable: MATCH				
	(1)	(2)	(3)	(4)	(5)
EDU	0.050***	0.049***	0.047***	0.048***	0.046***
	(12.90)	(12.65)	(12.43)	(12.44)	(12.09)
AGE	-0.0092***	-0.0092***	-0.0092***	-0.0092***	-0.0090***
	(-12.31)	(-12.23)	(-12.31)	(-12.31)	(-12.06)
IND(Reference: Services)					
[1] Agriculture	0.060	0.061	0.064	0.064	0.058
	(1.25)	(1.27)	(1.34)	(1.34)	(1.21)
[2] Energy	-0.185**	-0.187**	-0.195**	-0.194**	-0.198**
	(-2.74)	(-2.77)	(-2.89)	(-2.89)	(-2.95)
[3] Mining	-0.329*	-0.327*	-0.300	-0.300	-0.313
	(-2.05)	(-2.04)	(-1.84)	(-1.83)	(-1.94)
[4] Manufacturing	-0.059**	-0.055**	-0.052*	-0.052*	-0.053**
	(-2.81)	(-2.64)	(-2.51)	(-2.51)	(-2.53)
[5] Construction	0.146***	0.148***	0.147***	0.147***	0.145***
	(8.11)	(8.22)	(8.17)	(8.17)	(8.04)
[6] Trade	-0.137***	-0.136***	-0.132***	-0.132***	-0.135***
	(-7.07)	(-7.02)	(-6.83)	(-6.82)	(-6.98)
[7] Transport	-0.264***	-0.260***	-0.256***	-0.257***	-0.260***
	(-8.86)	(-8.69)	(-8.56)	(-8.57)	(-8.69)
[8] Bank,Insurance	0.020	0.017	0.014	0.014	0.013
	(0.54)	(0.45)	(0.36)	(0.37)	(0.34)
TOJCH(Reference: First job)					
Job After Break	-0.121***	-0.123***	-0.134***	-0.134***	-0.137***
	(-3.72)	(-3.80)	(-4.13)	(-4.13)	(-4.26)
Job With New Employer	-0.077*	-0.080**	-0.074*	-0.074*	-0.074*
	(-2.47)	(-2.58)	(-2.41)	(-2.41)	(-2.41)
Company Taken Over	0.200***	0.197***	0.179***	0.179***	0.168***
	(5.74)	(5.68)	(5.03)	(5.04)	(4.64)
Changed Job, Same Firm	0.030	0.024	-0.003	-0.003	-0.027
	(0.84)	(0.68)	(-0.07)	(-0.07)	(-0.72)
FSIZE(Reference: GE 2000)					
[1] LT 20	0.104***	0.104***	0.117***	0.117***	0.122***
	(5.14)	(5.16)	(5.77)	(5.77)	(6.00)
[2] GE 20 LT 200	0.037	0.038	0.044*	0.044*	0.049*
	(1.80)	(1.86)	(2.15)	(2.14)	(2.40)
[3] GE 200 LT 2000	-0.006	-0.006	-0.006	-0.006	-0.002
	(-0.29)	(-0.28)	(-0.28)	(-0.28)	(-0.07)
STATE(Reference: Bavaria)					
[1] Schleswig-Holstein	-0.043	-0.051	-0.047	-0.047	-0.047
	(-0.88)	(-1.04)	(-0.95)	(-0.96)	(-0.97)
[2] Hamburg	-0.007	-0.010	-0.015	-0.014	-0.017
	(-0.12)	(-0.16)	(-0.24)	(-0.23)	(-0.27)
[3] Lower Saxony	-0.004	-0.006	-0.007	-0.007	-0.008

Continued on next page

Variables	Dependent variable: MATCH				
	(1)	(2)	(3)	(4)	(5)
	(-0.14)	(-0.20)	(-0.21)	(-0.21)	(-0.26)
[4] Bremen	-0.067	-0.052	-0.050	-0.051	-0.048
	(-0.77)	(-0.60)	(-0.58)	(-0.59)	(-0.56)
[5] North-Rhine-Westfalia	0.005	0.009	0.010	0.010	0.009
	(0.19)	(0.34)	(0.37)	(0.38)	(0.33)
[6] Hessen	0.001	0.006	0.005	0.005	0.002
	(0.04)	(0.18)	(0.16)	(0.17)	(0.07)
[7] Rheinland-Pfalz	-0.019	-0.016	-0.016	-0.015	-0.017
	(-0.46)	(-0.39)	(-0.38)	(-0.38)	(-0.42)
[8] Baden-Wuerttemberg	0.016	0.025	0.022	0.023	0.022
	(0.56)	(0.87)	(0.78)	(0.78)	(0.75)
[10] Saarland	-0.108	-0.092	-0.091	-0.090	-0.094
	(-1.34)	(-1.15)	(-1.13)	(-1.12)	(-1.17)
[11] Berlin	-0.063	-0.061	-0.059	-0.059	-0.060
	(-1.42)	(-1.39)	(-1.34)	(-1.35)	(-1.37)
[12] Brandenburg	-0.096*	-0.112**	-0.112**	-0.112**	-0.110**
	(-2.29)	(-2.67)	(-2.68)	(-2.67)	(-2.63)
[13] Mecklenburg-Vorpommern	-0.094	-0.112*	-0.115*	-0.115*	-0.111*
	(-1.78)	(-2.12)	(-2.17)	(-2.17)	(-2.10)
[14] Saxony	0.013	-0.003	-0.004	-0.004	-0.003
	(0.40)	(-0.08)	(-0.13)	(-0.12)	(-0.09)
[15] Saxony-Anhalt	-0.058	-0.076	-0.074	-0.074	-0.070
	(-1.38)	(-1.79)	(-1.76)	(-1.75)	(-1.66)
[16] Thuringia	-0.065	-0.085*	-0.079	-0.079	-0.077
	(-1.56)	(-2.02)	(-1.89)	(-1.89)	(-1.84)
Survey year t (Reference: 2000)					
2001	0.028	0.028	0.027	0.027	0.028
	(0.97)	(0.96)	(0.95)	(0.95)	(0.97)
2002	0.041	0.040	0.038	0.038	0.034
	(1.35)	(1.31)	(1.25)	(1.24)	(1.13)
2003	0.018	0.016	0.018	0.018	0.016
	(0.58)	(0.52)	(0.57)	(0.56)	(0.50)
2004	0.025	0.025	0.026	0.026	0.021
	(0.80)	(0.79)	(0.81)	(0.82)	(0.66)
2005	0.053	0.052	0.053	0.053	0.048
	(1.67)	(1.62)	(1.67)	(1.67)	(1.50)
2006	0.015	0.014	0.015	0.014	0.009
	(0.47)	(0.44)	(0.46)	(0.46)	(0.30)
2007	0.013	0.013	0.020	0.019	0.013
	(0.44)	(0.43)	(0.64)	(0.63)	(0.41)
2008	0.056	0.056	0.061*	0.060*	0.054
	(1.84)	(1.85)	(2.00)	(1.99)	(1.77)
2009	0.082**	0.084**	0.086**	0.086**	0.079**
	(2.72)	(2.76)	(2.85)	(2.85)	(2.61)
2010	0.036	0.037	0.027	0.027	0.017
	(1.11)	(1.13)	(0.83)	(0.83)	(0.51)
2011	0.072*	0.072*	0.065*	0.065*	0.052
	(2.20)	(2.20)	(1.98)	(1.97)	(1.58)
2012	0.104**	0.103**	0.099**	0.099**	0.087**
	(3.13)	(3.09)	(2.95)	(2.94)	(2.58)
2013	0.105**	0.105**	0.100**	0.100*	0.088**
	(3.15)	(3.13)	(2.97)	(2.97)	(2.62)
2014	0.126***	0.133***	0.130***	0.130***	0.116***
	(4.24)	(4.50)	(4.37)	(4.37)	(3.89)
ISEI	0.0107***	0.0105***	0.0102***	0.0102***	0.0100***

Continued on next page

Variables	Dependent variable: MATCH				
	(1)	(2)	(3)	(4)	(5)
	(18.99)	(18.62)	(18.21)	(18.20)	(17.90)
MIGBACK (Reference: German national)					
Direct migrant		-0.142***	-0.132***	-0.126***	-0.129***
		(-5.21)	(-4.87)	(-3.86)	(-4.74)
Indirect migrant		-0.051	-0.049	-0.059	-0.049
		(-1.54)	(-1.51)	(-1.59)	(-1.50)
Referrals			-0.138***	-0.138***	
			(-9.71)	(-8.90)	
MIGBACK × REF (Reference: German national × Formal channels)					
Direct migrant × referrals				-0.016	
				(-0.35)	
Indirect migrant × referrals				0.029	
				(0.56)	
Chan (Reference: Internet)					
Public emp. agency					-0.107***
					(-3.61)
Private emp. agency					-0.075
					(-1.34)
Newspaper					-0.075**
					(-2.74)
Referrals					-0.175***
					(-7.16)
Other					-0.004
					(-0.17)
ρ	0.642	0.642	0.639	0.639	0.637
Std. Err. of ρ	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
χ^2	1307.16	1301.92	1280.71	1281.12	1272.18
Prob$\geq \chi^2$	0.00	0.00	0.00	0.00	0.00
LR test(Prob$> \chi^2$)		0.00	0.00	0.7885	0.00
Observations	15779	15779	15779	15779	15779
Pseudo R^2	0.178	0.180	0.184	0.184	0.186

t statistics in parentheses. Marginal effects for factor levels is the discrete change from the base level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Chapter 4

Search Channels of Workers and Firms over the Business Cycle

4.1 Introduction

This study investigates the search channels of workers and firms over the business cycle. In particular, how the firms and the workers change the intensity of job advertisement and the job search through different search channels during expansions or recessions. Let us observe the case of expansion. During expansions when there are many vacancies and few unemployed workers it is relatively easier for workers to find jobs, so we would expect the workers to rely less strongly on their social networks. While the firms have difficulties to fill their open positions in expansions and rely more strongly on their social networks, which means that referral hiring may be dominating in expansions. As a result, there are two counteracting effects of expansion on the proportion of referral hiring. The main objective of this paper is to find out which of the effects dominates.

Empirical analysis are conducted using the data from the German Socio-Economic Panel (SOEP) over the period 2000-2014 and the data from the IAB Job Vacancy Survey over the period 2000-2013. Further, this study uses the data on the annual GDPs of 16 German federal states provided by the Federal Statistical Office of Germany. Using the two-step estimation procedure suggested by Solon et al. (1994) we show that in the long run there is a positive correlation between the GDP and the proportion of workers hired through their social networks.

In order to explain the effect of the productivity change on the search and matching strategies of firms and workers, this paper presents a search and matching model with two search channels. In the model workers and firms are matched either through referrals, or through formal search channels. The firm chooses advertisement effort through formal search channels to maximize the asset value of an open vacancy. Workers choose their search intensity through formal channels to maximize the asset value of being unemployed. The wages are determined using the Nash bargaining rule, where an individual firm and worker do not influence the behavior in the rest of the labour market.

Calibration results of the model show that during expansions unemployment rate is decreasing, vacancy rate and wages are increasing. As a result, employment becomes more gainful for workers, and they increase the job search effort through formal channels. This behaviour of workers reduces the proportion of referral hiring. While firms react differently. On the one hand the productivity increases, so the advertisement of the vacancy through formal channels should become more gainful for the firms. On the other hand, there are less unemployed workers available per vacancy issued, and wages are higher during expansions. Thus, the advertisement of the vacancy through formal channels should become less gainful for the firms. According to the calibration results advertisement effort of the firms through formal channels decreases during expansion, which means that the negative effect on the job advertisement dominates. Since the advertisement effort of the firms in the formal search channels decreases, intuitively, we would expect that the fraction of workers hired through referrals should increase. Calibrations show that the proportion of workers hired through referrals increases during expansions. Thus the negative effect of the advertisement effort decrease on the fraction dominates the positive effect of the search effort increase. Both the estimation and the calibration results indicate that the firm-side effect dominates.

There is a large literature about the job search methods used by workers and recruitment methods used by the firms. One of the earlier studies by Granovetter (1974) emphasizes the importance of the social contacts when observing search channels of the firms and workers. In the literature informal search methods of firms and workers include the following search channels: search through relatives, friends, acquaintances, referrals from other employees and etc. Formal search methods include search through newspaper advertisements and advertisements in the internet, search through the state and private employment agencies, school and college placement services and etc.

Different theoretical and empirical studies were conducted to show the effect of search methods of firms and workers on the labour market outcomes. Particularly, there is a strand of literature which study the effect of using referrals on wages and the probability of being hired. Montgomery (1991), Kugler (2003), Dustmann et al. (2016) and Galenianos (2013) find positive wage effect of using referrals. While Pistaferri (1999), Bentolila et al. (2010) and Zaharieva (2018) find the effect to be negative.

Montgomery (1991) finds that high ability workers will refer to an own type high ability workers based on the fact that social contacts tend to occur among workers with similar characteristics. Moreover, a worker will refer only well-qualified applicants, since his reputation is at stake. Thus the employer tries to mitigate the adverse-selection problem and the workers hired through referrals are paid higher wages. According to Kugler (2003) referees exert peer pressure on the referred workers, which enables the employer to reduce monitoring cost, and to pay lower efficiency wages to the workers hired through referrals. So the workers with relatively more contacts are matched to the jobs with higher wages. While Dustmann et al. (2016) and Galenianos (2013) assume that the the worker's match-specific productivity is less uncertain when using informal search methods compared to formal search methods. As a result both authors find that the match quality is better in case of hiring through referrals compared to the hiring through formal search channels. As a result, the workers hired through informal search methods initially get higher wages. Whereas Bentolila et al. (2010) and Zaharieva (2018) argue that referrals are more likely to cause a mismatch between worker's most productive and actual occupation. Thus the use of informal methods by the workers leads to lower quality matches between workers and firms and results in a wage penalty.

While these studies observe the effect of using referrals on the labour market outcomes, this study rather aims to find the effect of labour market conditions on the relative frequency of using referrals. In this sense, the present study is closely related to the study by Galeotti and Merlino (2003). The author finds that the use of networks and their efficiency in matching workers to jobs depends on the labor market conditions. They assume that the workers invest in their networks to keep the network active. Further, incentives in networking relate to labour market conditions, in particular, to the separation rate. When the separation rate is low, the risk of becoming unemployed is low, so is the investment of workers in their networks. When the separation rate is high, then the workers connected to a job seeker are more likely searching for job themselves, moreover, they have more contacts who they can pass the information about a vacancy. As a result, investment in networks is less gainful, and workers' investment in networks is low. Workers' investment in networks is high, when the separation rates are moderate. Thus, the relation between the job separation rate and the probability of a worker to find a job through referrals is inverse U-shaped.

Another paper by Shimer (2004) finds that when the job search cost is low, the response of the workers' search intensity to labour market condition may be acyclical. The intuition is that during expansions even low job search effort may be enough for the worker to find a job. He finds that when the labour market conditions are weak, the workers who have high probability to get hired, increase their search intensity, while others get discouraged and reduce their search intensity. Unlike Shimer (2004), our model predicts that the search intensity of workers is procyclical. This result is inline with the findings of Pissarides (2000).

The study proceeds as follows: section 2 describes the data used, discusses the empirical approach and estimation results; section 3 explains the search and matching model with several search channels, where the firms and the workers choose the intensity of advertisement and job search respectively; section 4 presents the results from the calibration of the model; section 5 concludes.

4.2 Empirical Analysis

4.2.1 Data

In this section we empirically estimate the effect of the productivity change on the search channels of the workers and the firms. Particularly, we check if the change of the GDP has significant effect on the proportion of the referral hiring, and if yes, whether the effect is positive of negative.

This subsection describes the two datasets used for the empirical analysis. This study uses data from the German Socio-Economic Panel (SOEP), and the IAB Job Vacancy Survey of the Institute for Employment Research (IAB JVS). The German Socio-Economic Panel is a longitudinal study of households and individuals, which annually covers about 11,000 households, and 30,000 individuals. The surveyed individuals are asked a wide range of questions regarding personal characteristics and employment data. In addition, currently employed respondents are asked how they found their current job, the possible answers being, through federal employment office, an advertisement in the internet or newspaper, a job-center (ARGE) and a private recruitment agency, other, and through relatives and friends. The observations where workers found the job through relatives and friends are treated to be caused by referral hiring. The rest of the search channels are treated as a formal search channels in the analysis. Our sample includes 19148 observations on employed individuals from SOEP 2000-2014.

Whereas, IAB JVS is an annual cross-sectional survey of firms, which is representative for the German firms. The sample of the firms is drawn yearly from all firm size categories, all the industries and regions of Germany. Among a wide range of questions regarding the characteristics of the firms, the human resource managers or managing directors answer to questions about the most recent recruitment case. In particular, they answer about the recruitment channel which led to hiring of the last hired employee. In this study the recruitment channels are classified into formal channels and referrals as follows. The employee is hired through referrals when the recruitment takes place via own employees/personal contacts, or through internal job advertisement. The rest of the recruitment channels are included into the formal channels category. The formal recruitment channels include recruitment through private placement service, Federal Employment Agency, advertisements in newspapers and magazines, posting the vacancy to internet website, choosing from unsolicited applications/pool of applicants and etc. During the survey 7500 to 15000 establishments are surveyed annually. Our sample of the IAB JVS includes 59329 observations on recruitment cases from 2000-2013.

As a measure of the productivity annual GDPs of 16 German federal states are used in the estimations. The data is provided by the Federal Statistical Office of Germany includes unadjusted regional GDP at current prices in Euro.

4.2.2 Empirical Approach

To estimate the effect of GDP on the proportion of the referral hiring we use the two-step estimation procedure suggested by Solon et al. (1994). The estimation approach is used to deal with the Moulton (1986) problem. According to Moulton (1986) individuals in the same year have a component of variance in common, that can be entirely attributed neither to the individual characteristics, nor to the aggregate variables in the year. As a result the standard error of the aggregate variables is underestimated, because of the positive correlation of the error component across the individuals.

To get around with the problem on the first step $REF_{j,i,t}$ is regressed on the control variables, state and time dummies. $REF_{j,i,t}$ equals 1, if the j^{th} recruitment is done through referrals, at time t, and in the state i. $REF_{j,i,t}$ takes value 0, if the j^{th} recruitment is done through formal channels, at time t, and in the state i. In case of using the SOEP data j is the identifier for the individual, while in case of using the JVS data j is the identifier for the firm. The first step of the estimation is characterized by the following regression equation:

$$REF_{j,i,t} = B_0 + \mathbf{BX} + \sum\sum s_{i,t}D_{i,t} + \gamma_{j,i,t} \tag{4.1}$$

where $\mathbf{B}$ is the vector of coefficients on the vector of exogenous variables $\mathbf{X}$. $D_{i,t}$ are the time and state dummies. $\gamma_{i,t}$ are the i.i.d. error terms.

When using SOEP data several exogenous variables are used on the first step of the estimations. Among them there are variables indicating the individuals' education[1], age, gender and the nationality. In addition we include a categorical variable with four categories showing the size of the firm in which the j^{th} individual is employed at time t. $FSIZE_{j,i,t}$ is a categorical variable with four categories showing the size of the firm in which the j^{th} individual is employed at time t. The four categories are: less than 20 employees, 20 to 200, 200 to 2000, and more than 2000 employees. Another categorical variable included indicates the industry of the firm where the individual is employed at time t. The variable has 9 categories: Agriculture, Energy, Mining, Manufacturing, Construction, Trade, Transport, Bank/Insurance, and Services. Yet another categorical variable included has 5 categories and indicates which kind of job change preceded the current employment of the individual. The categories of the exogenous variable are the following: first job, job after break, job with new employer, company taken over, changed job at the same firm. Finally, the Standard International Socio-Economic Index of Occupational Status developed by Ganzeboom et al. (1992) is included as a control for the occupational status. ISEI index reflects individual's socio-economic status based on information about the individual's income, education, and occupation. It takes values in the range between 16 and 90.

When the JVS data is used on the first step, the control variables include age and gender of the last hired employee. Moreover, we control for the previous employment status of the employee. The corresponding categorical variable has 9 categories with information if previously the employee was unemployed, employed elsewhere, self-employed, in vocational training/further education, temporary worker at the establishment and etc. A categorical variable with 6 categories shows the size of the firm. The six categories are: less than 10 employees, 10 to 19, 20 to 49, 50 to 199, 200 to 499 and more than 499 employees. Additionally we control for the search duration of the firm using the variable indicating

[1]As an exogenous variable measuring the individual's education we use the amount of education or training in years computed by the SOEP (for detailed description see Helberger (1988) and Schwarze et al. (1991). The corresponding variable ranges from 7 to 18.

the search duration in days. Further, two binary choice variables are used the one indicating if there were impediments or external reasons during the past 12 months, which prevented the firm to fully use its opportunities, the second, indicating whether the firm searched for employees in vain during the past 12 months.

On the second step the time and state fixed effects $\hat{s_{i,t}}$ estimated on the first step are regressed on the GDP, the time and state dummies, and the interaction term between the state GDPs and state dummies. The interaction term is included in the regression equation to control for the regional difference of the effect of state GDP. Thus, the second step of the estimation is characterized by the following regression equation:

$$\hat{s_{i,t}} = a_i + \beta_1 GDP_{i,t} + \beta_{2i} GDP_{i,t} \times a_i + \delta_t + \tau_{i,t} \tag{4.2}$$

where δ_t is the time fixed effect, a_i is the state fixed effect, and $\tau_{i,t}$ are the i.i.d. error terms. Thus, the sign of the sum $\beta_1 + \beta_{2i}$ determines the sign of the effect of state specific GDPs on the proportion of workers recruited through referrals at the states. The estimation results of the second step using the SOEP and JVS data are summarized in the Table 4.3 and the Table 4.4 respectively.

4.2.3 Descriptive Statistics

This section presents descriptive statistics of the variables used in the estimations. Table 4.1 reveals descriptive statistics of control variables from SOEP data. Means of continuous variables and percentages of observations in each category for categorical variables are displayed for the subgroups of individuals which found job through referrals, formal channels and total. SOEP data demonstrates that 32.43% of workers found their job through referrals. Further, those who found job through referrals are younger, less educated, and have lower ISEI socio-economic status than those workers which found their job through formal channels. Proportionally there are less females and more immigrants among those which found job through referrals. They are proportionally more frequently working in Agriculture, Mining, Manufacturing, Construction, Trade, and Transport, while less frequently working in Energy, Bank/Insurance, and Services. Finally, the last job change of those which found job through referrals is proportionally more often categorized as "First Job", "Job With New Employer", and less often as "Job After Break", "Company Taken Over" and "Change Job, Same Firm".

Table 4.2 presents descriptive statistics of control variables from JVS data. The data reveals that on average 32.72% of workers are recruited through referrals. Unlike SOEP data, JVS data shows that workers hired through referrals are older and proportionally more often they are females compared to workers hired through formal channels. But both datasets show that workers hired through referrals are working at smaller companies. Previous employment status of workers hired through referrals is proportionally more often categorized as "Employed elsewhere", "Self-employed", "Did not work (e.g. housewife)", and "Other", and less often as "Unemployed", "Temporary worker ate the establishment", and in vocational training/further education at the establishment or elsewhere. Search duration is shorter for the workers recruited through referrals. Further, firms which hired through referrals experienced more impediments or external reasons during the past 12 months, which prevented the firm to fully use its opportunities, and they searched for employees in vain less often during the past 12 months.

Table 4.1: Descriptive statistics of control variables.

	SOEP		
Variable	Referrals	Formal channels	Overall
Education	12.02	12.75	12.51
Age	33.85	34.39	34.21
Gender			
Male	46.02	44.63	45.08
Female	53.98	55.37	54.92
Nationality			
Germans	89.57%	93.58%	92.28%
Immigrants	10.43%	6.42%	7.72%
Firm size			
LT 20	42.62%	29.68%	33.88%
GE 20 LT 200	29.40%	29.66%	29.58%
GE 200 LT 2000	14.43%	19.11%	17.59%
GE 2000	13.54%	21.54%	18.95%
Industry			
Agriculture	1.74%	1.51%	1.58%
Energy	0.52%	1.02%	0.86%
Mining	0.16%	0.12%	0.14%
Manufacturing	14.90%	13.93%	14.24%
Construction	11.96%	11.66%	11.76%
Trade	22.98%	17.04%	18.97%
Transport	5.94%	5.51%	5.65%
Bank/Insurance	2.24%	3.86%	3.34%
Services	39.57%	45.35%	43.47%
Type of job change			
First Job	12.77%	11.10 %	11.64%
Job After Break	24.06%	26.55%	25.74%
Job With New Employer	60.43%	46.91%	51.30%
Company Taken Over	0.89%	3.80%	2.86%
Changed Job, Same Firm	1.85%	11.64%	8.47%
ISEI	40.83	45.87	44.23
Observations	6210	12938	19148
	32.43%	67.57%	100%

Before discussing estimation results, let us explore aggregated data for the study period from 2000 to 2014. Figure 4.1. presents aggregated descriptive statistics on the proportion of referral hiring and GDP of Germany over time.

Table 4.2: Descriptive statistics of control variables.

JVS			
Variable	Referrals	Formal	Overall
Age*	36.14	34.88	35.29
Gender			
Male	41.00%	47.25%	45.20%
Female	59.00%	52.75%	54.80%
Search duration†	49.44	58.61	55.79
Firm size			
LT 10	20.12%	11.07%	14.03%
GE 10 LT 20	22.59%	16.49%	18.49%
GE 20 LT 50	28.25%	25.28%	26.25%
GE 50 LT 200	18.00%	24.79%	22.57%
GE 200 LT 500	6.07%	10.94%	9.34%
GE 500	4.98%	11.43%	9.32%
Previous employment status †			
Unemployed	34.44%	36.97%	36.14%
Employed elsewhere	45.37%	42.08%	43.16%
Self-employed	3.42%	2.22%	2.61%
Vocational training/further education	7.18%	8.60%	8.13%
Did not work (e.g. housewife)	3.71%	2.37%	2.81%
Temporary worker at the establishment	1.87%	2.72%	2.45%
Vocational training at the establishment	1.69%	2.07%	1.95%
Vocational training/further education elsewhere	2.10%	2.85%	2.60%
Other	0.22%	0.11%	0.15%
Impediments or external reasons			
Yes	34.80%	32.04%	32.94%
No	62.53%	65.25%	64.36%
Not specified	2.68%	2.71%	2.70%
Search in vain			
Yes	17.27%	20.92%	19.72%
No	76.82%	73.79%	74.78%
Not specified	5.91%	5.30%	5.50%
Observations	29556	60777	90333
	32.72%	67.28%	100%

* Age is available for 88560 individuals.

‡ Search duration is available for 72341 individuals.

† Survey methodology of previous employment status changed during the study period.From 2000 to 2003 respondent firms could choose options: Unemployed, Employed elsewhere, Vocational training/further education, Did not work (e.g. housewife), and Other. Starting from 2004 to 2007 Self-employed was added to the list. From 2008 to 2010 another option Temporary worker at the establishment was added. Throughout 2011 to 2013 the option Vocational training/further education was split to two different answers: Vocational training at the establishment and Vocational training/further education elsewhere. 90327 responses were recorded determining previous employment status in our dataset.

The left y-axis of the two y-axis figure shows the proportion of referral hiring in per-

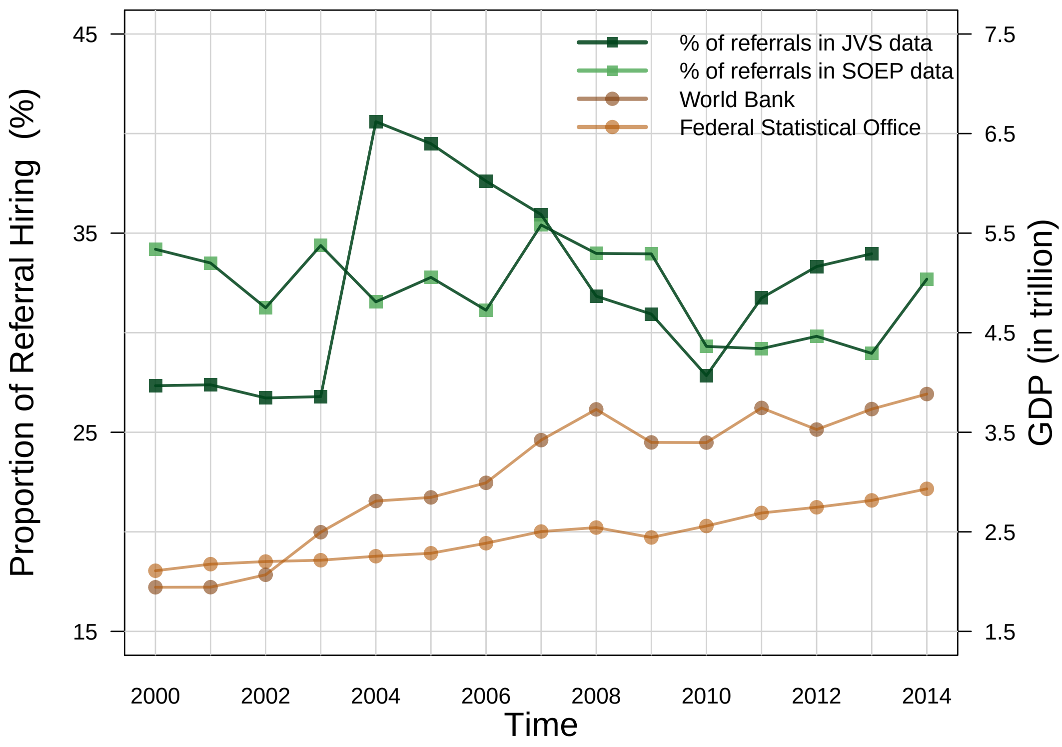

Figure 4.1: Proportion of referral hiring and GDP over time.

centage observed in SOEP and JVS datasets. The right y-axis shows the GDP of Germany measured in trillion. In the data from the World Bank GDP is measured at current prices in US dollars. The calendar and seasonally adjusted measure from the Federal Statistical Office of Germany is in euros. No clear pattern of correlation between the proportion of referral hiring and GDP can be determined in the figure. But we can see that proportion of referral hiring is decreasing in the period from 2008 to 2010, and the it starts recovering in both SOEP and JVS data.

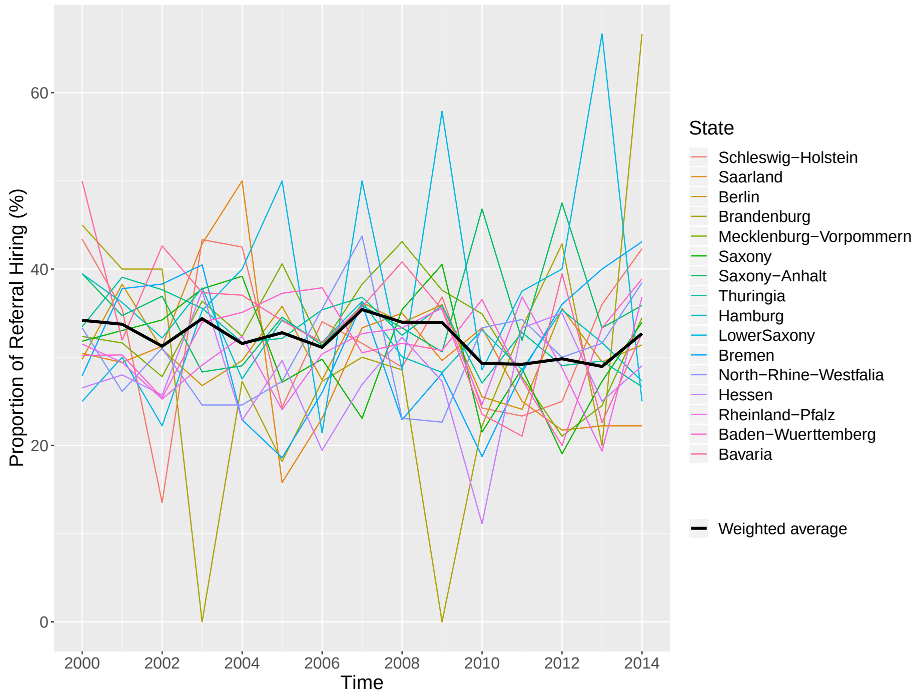

Figure 4.2: Regional proportions of referral hiring over time.

On the next step we observe the proportion of referral hiring in SOEP data by plotting the regional proportions, and average over time in Figure 4.2. The average is weighted by the number of observations from the region. We can observe large variability of proportions among the states, and different trajectories of regional proportions over time. These observations outline the importance of estimating regional proportions instead of proportion at country level.

Regional differences are also observed with respect to GDPs. Figure 4.3 depicts GDPs of German federal states over time. The calendar and seasonally adjusted data is provided by the Federal Statistical Office of Germany. Nominal values of state GDPs and their natural logarithms are presented in Figure 4.3 and Figure 4.4 respectively. Although regional GDPs follow similar pattern over time, there are large differences in their absolute values. This further motivates estimating regional proportions instead of proportion at country level. By estimating regional proportions, first, we are able to control for the state fixed effects and regional differences in the effect of state GDP. Second, larger data is exploited for statistical inference on the second step of the statistical approach, improving statistical significance of estimation results.

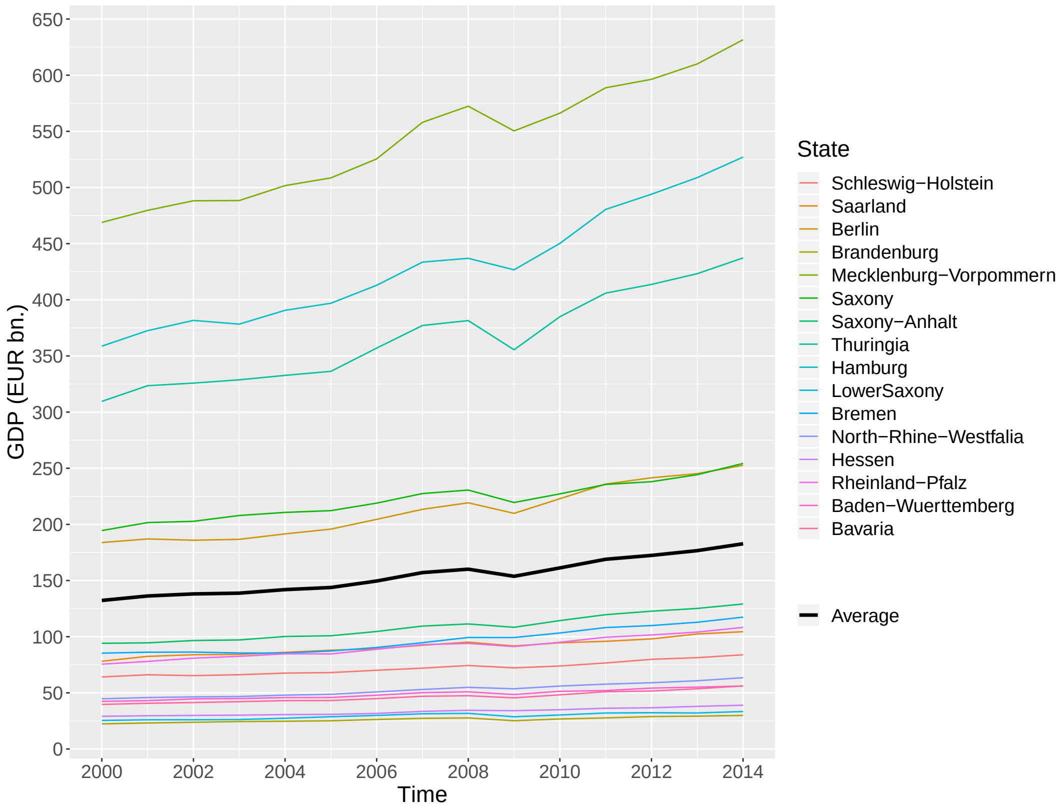

Figure 4.3: Nominal values of state GDPs over time.

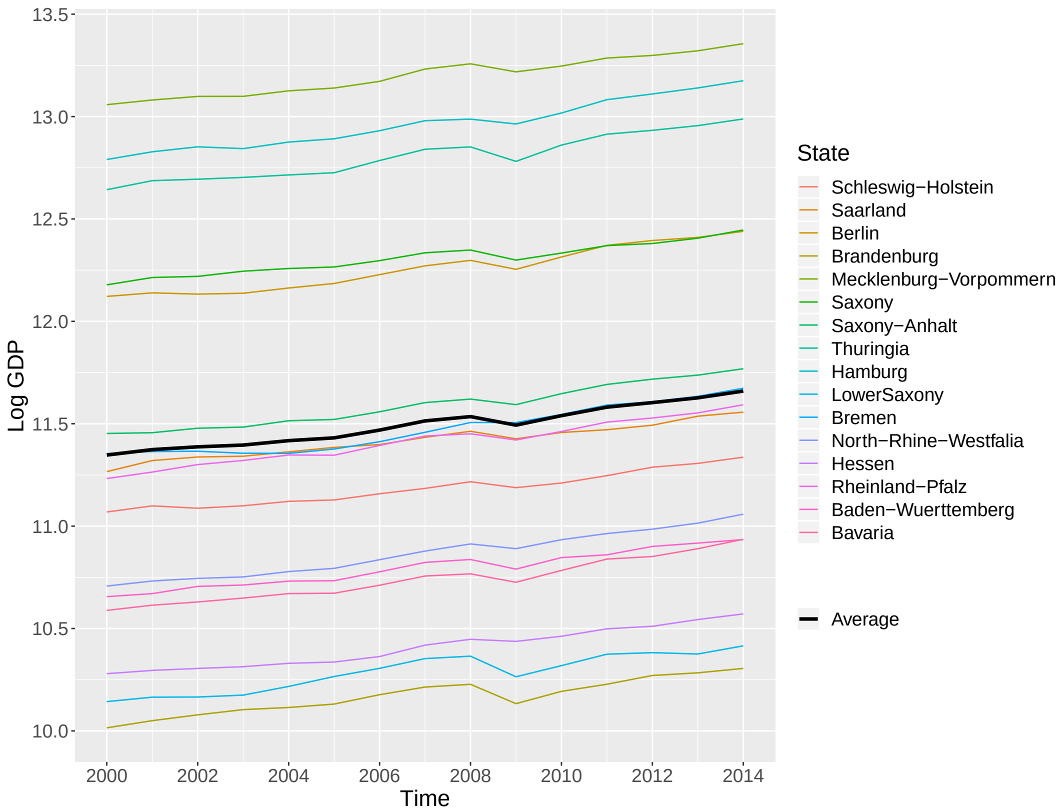

Figure 4.4: Natural logarithms of state GDPs over time.

4.2.4 Estimation Results

The first column of Table 4.3 presents the results of the estimations of the second step using state and time fixed effects and state specific effects of the natural logarithms of GDPs. According to the results the effect of GDP on the proportion of referrals used in recruitment $\beta_1 + \beta_{2i}$ is positive for all the states. Moreover, the coefficients of the state specific effects are statistically significantly different from the coefficient of the reference state Bavaria in case of all the states except Baden-Wuerttemberg. The significant differences highlight the importance of using interaction terms. Further, in the column 2 the time trend is used instead of the time fixed effects. In this case the effects are positive for all the states except Hamburg and Thuringia. The time trend turns to be negative and statistically significant, meaning that the proportion of referrals used in recruitment is decreasing over the time. It is important to note that compared to the first regression equation, the adjusted R squared is smaller in the second regression equation.

When the nominal levels of GDPs are used instead of their natural logarithms in the columns 3 and 4, the results are similar. In case of using time fixed effects all the coefficients $\beta_1+\beta_{2i}$ are again positive. The more, coefficients of the state specific effects are statistically significantly different from the coefficient of the reference state in case of all the states. While when using the time trend as an explanatory variable the state specific effects for four states become negative. And again the time trend is negative and statistically significant. Compared to the third regression equation, the adjusted R squared is smaller in the fourth regression equation indicating higher explanatory power of time fixed effects compared to the time trend. The empirical results obtained using the SOEP data suggest that the effect of GDP on the proportion of workers hired through referrals is positive.

When the time and state fixed effects $\hat{s_{i,t}}$ from the first step of estimations using the JVS data are used as a dependant variable the coefficients obtained are qualitatively similar. The column 1 in Table 4.4 shows that the effects of the state specific natural logarithms of GDPs are positive in all the states except Bremen and Berlin. Furthermore, the coefficients are significantly different from the coefficient of Bavaria in all the states except North-Rhein-Westfalia and Saarland, again indicating that the interaction terms should be included into the regression equation. In the regression equation represented by the column 2 the time fixed are substituted with a time trend. The trend is negative and statistically significant. Additionally, one more state specific effect of $logGDP$ in the Saxony becomes negative compared to the column 1. The adjusted R squared is lower for the regression equation of the column 2 compared to the one of the column 1.

In the columns 3 and 4 the nominal levels of GDPs are used instead of their natural logarithms as explanatory variables. When using time fixed effects in the regression equation all the coefficients $\beta_1 + \beta_{2i}$ are positive except those in Bremen and Berlin. While the coefficients of the state specific effects are significantly different from the coefficient of Bavaria in case of all the states except Saxony and Saxony-Anhalt. Whereas, in the column 4 the time fixed effects are substituted with the time trend, and negative coefficients are

Table 4.3: Estimation results of the second step using SOEP data

	(1) T&S FE ($s_{i,t}$)	(2) T&S FE ($s_{i,t}$)	(3) T&S FE ($s_{i,t}$)	(4) T&S FE ($s_{i,t}$)
Reference year: 2000	$logGDP_{i,t} \times State$		$GDP_{i,t} \times State$	
Reference state: Bavaria	0.512***	0.069***	0.00123***	-0.00005
	(0.027)	(0.018)	(0.00006)	(0.00004)
Schleswig-Holstein	0.912***	0.290***	0.01316***	0.00238***
	(0.028)	(0.029)	(0.00056)	(0.00044)
Hamburg	0.621**	-0.022*	0.00718***	-0.00177***
	(0.037)	(0.040)	(0.00052)	(0.00046)
Lower Saxony	0.683***	0.189***	0.00332***	0.00041***
	(0.013)	(0.014)	(0.00009)	(0.00007)
Bremen	1.019***	0.487***	0.04251***	0.01625***
	(0.045)	(0.049)	(0.00210)	(0.00201)
North-Rhine-Westfalia	0.817***	0.300***	0.00155***	0.00033***
	(0.013)	(0.013)	(0.00002)	(0.00002)
Hessen	0.829***	0.179***	0.00391***	0.00019*
	(0.022)	(0.021)	(0.00014)	(0.00011)
Rheinland-Pfalz	0.793***	0.281***	0.00751***	0.00162***
	(0.017)	(0.018)	(0.00025)	(0.00019)
Baden-Wuerttemberg	0.534	0.049	0.00151***	-0.00015**
	(0.012)	(0.013)	(0.00003)	(0.00003)
Saarland	0.866***	0.441***	0.02994***	0.01114***
	(0.037)	(0.041)	(0.00158)	(0.00150)
Berlin	0.882***	0.399***	0.00937***	0.00323***
	(0.017)	(0.019)	(0.00027)	(0.00021)
Brandenburg	0.847***	0.388***	0.01663***	0.00555***
	(0.017)	(0.019)	(0.00056)	(0.00042)
Mecklenburg-Vorpommern	0.809***	0.275***	0.02534***	0.00502***
	(0.025)	(0.027)	(0.00115)	(0.00093)
Saxony	0.775***	0.312***	0.00893***	0.00236***
	(0.014)	(0.016)	(0.00029)	(0.00021)
Saxony-Anhalt	0.832***	0.266***	0.01765***	0.00293***
	(0.022)	(0.023)	(0.00076)	(0.00056)
Thuringia	0.366***	-0.141***	0.00856***	-0.00507***
	(0.018)	(0.020)	(0.00068)	(0.00050)
Time trend		-0.00577***		-0.00338***
		(0.00043)		(0.00039)
Time FE	v		v	
State FE	v	v	v	v
Constant	-6.602***	10.615***	-0.49679***	6.73418***
	(0.349)	(0.660)	(0.36948)	(0.76842)
Observations	19148	19148	19148	19148
Adjusted R^2	0.2988	0.1273	0.3016	0.1247

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4.4: Estimation results of the second step using JVS data

	(1)	(2)	(3)	(4)
	T&S FE ($s_{i,t}$)	T&S FE ($s_{i,t}$)	T&S FE ($s_{i,t}$)	T&S FE ($s_{i,t}$)
Reference year: 2000	$logGDP_{i,t} \times State$		$GDP_{i,t} \times State$	
Reference state: Bavaria	0.147***	0.083***	0.00031***	0.000002
	(0.014)	(0.012)	(0.00003)	(0.00003)
Schleswig-Holstein	0.219***	0.109	0.00268***	-0.00018
	(0.018)	(0.024)	(0.00033)	(0.00036)
Hamburg	0.321***	0.260***	0.00327***	0.00145***
	(0.021)	(0.028)	(0.00029)	(0.00033)
Lower Saxony	0.659***	0.575***	0.00292***	0.00218***
	(0.009)	(0.012)	(0.00005)	(0.00006)
Bremen	-0.184***	-0.211***	-0.00785***	-0.01226***
	(0.030)	(0.041)	(0.00132)	(0.00164)
North-Rhine-Westfalia	0.159	0.090	0.00025***	-0.00004
	(0.009)	(0.011)	(0.00002)	(0.00002)
Hessen	0.185*	0.155***	0.00070***	0.00006
	(0.016)	(0.019)	(0.00009)	(0.00009)
Rheinland-Pfalz	0.359***	0.281***	0.00311***	0.00167***
	(0.012)	(0.017)	(0.00015)	(0.00016)
Baden-Wuerttemberg	0.114***	0.037***	0.00025*	-0.00016***
	(0.008)	(0.011)	(0.00002)	(0.00003)
Saarland	0.149	0.199***	0.00427***	0.00317*
	(0.026)	(0.036)	(0.00102)	(0.00127)
Berlin	-0.231***	-0.402***	-0.00248***	-0.00487***
	(0.009)	(0.013)	(0.00013)	(0.00013)
Brandenburg	0.087***	0.033***	0.00129***	-0.00123***
	(0.009)	(0.012)	(0.00028)	(0.00026)
Mecklenburg-Vorpommern	0.211***	0.112	0.00565***	0.00009
	(0.011)	(0.015)	(0.00055)	(0.00052)
Saxony	0.022***	-0.010***	0.00004	-0.00122***
	(0.007)	(0.010)	(0.00015)	(0.00013)
Saxony-Anhalt	0.049***	0.009***	0.00055	-0.00213***
	(0.011)	(0.014)	(0.00038)	(0.00034)
Thuringia	0.192***	0.147***	0.00372***	0.00098***
	(0.009)	(0.012)	(0.00035)	(0.00031)
Time trend		-0.00054*		0.00172***
		(0.00027)		(0.00026)
Time FE	v		v	
State FE	v	v	v	v
Constant	-1.875***	0.062	-0.10841***	-3.39681***
	(0.182)	(0.425)	(0.01201)	(0.50644)
Observations	59329	59329	59329	59329
Adjusted R^2	0.5751	0.1935	0.5723	0.1898

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

obtained for the half of the states. But in the latter case the adjusted R squared is much lower compared to the statistic in the column 3.

To sum up, the effect of the measure of productivity on the proportion of workers hired through referrals is positive in most of the states. The same is true in the estimations using both SOEP and JVS data. Moreover, in all the regression equations most of the coefficients of the interaction terms are significantly different from the reference state Bavaria. Which shows that the interaction terms should be included into the regression equation on the second step of the estimation. According to our empirical findings during the expansions when unemployment is relatively lower and there are relatively more vacancies, the proportion of referral hiring is higher. We propose a theoretical model to better understand the causes of the changes. In particular we explore the changes of the firms' advertisement intensity through the formal channels and the search effort of the workers through the formal channels over the business cycle.

4.3 Model

We develop a search and matching model with several search channels using 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 real productivity of the workers is p, and each worker has h number of contacts. The workers can be matched to the firm through referrals or through formal channels.

4.3.1 Matching function

The matching functions for the two matching channels are different. Let u be the unemployment rate of workers, and s be the search intensity through formal channels of the representative unemployed worker. su can be defined then as the efficiency units of searching workers. Furthermore, let v be the vacancy rate, and a be the level of job advertising of the representative hiring firm through formal channels. Thus, av can be defines as the efficiency units of job vacancies. We assume that the matching function through formal channels $m_f(su, av)$ has Cobb-Douglas form and write it as

$$m_f(su, av) = \lambda_0 (su)^{1-\eta} (av)^{\eta} \tag{4.3}$$

Where λ_0 is the total factor productivity of the matching function, and η is the elasticity with respect to the vacancy in the matching function. Note that s and a are the market averages. Given the search intensity of the representative worker and the job advertising of the representative firm the individual workers and firms choose their own search intensities and job advertising levels respectively. Let s_j be the search intensity of the worker j. Unemployed workers move from unemployment into employment according to a Poisson process with rate $m_f(su, av)/su$. The job finding rate of the worker j in the formal market can be written as

$$\lambda_{fj} = \frac{s_j}{su} m_f(su, av) \tag{4.4}$$

The job filling rate in the formal market of the firm is derived as follows. Let a_i be the job advertising level of the firm i. A vacant job is filled according to a Poisson process with rate $m_f(su, av)/av$. So the job filling rate of the firm i in the formal market can be written as

$$q_{fi} = \frac{a_i}{av} m_f(su, av) \tag{4.5}$$

In the symmetric Nash equilibrium all workers choose the same search intensity $s_j = s$, and all firms choose the same job advertising level $a_i = a$. Thus equations describing the job finding rate and the job filling rate for the representative worker and representative firm respectively can be written as follows:

$$\lambda_f = \frac{m_f(su, av)}{u} \tag{4.6}$$

$$q_f = \frac{m_f(su, av)}{v} \tag{4.7}$$

Next, we explain the mechanism of referral hiring. Following Stupnytska and Zaharieva (2015), matching function in case of matching through referrals m_r can be written as:

$$m_r = v\alpha((1-u)(1-(1-u)^h) \tag{4.8}$$

Where α is the exogenous rate at which a vacancy arrives to a worker per unit time. (see Cahuc and Fontaine (2009).) The firm and the workers are matched through referrals as follows. The firm opens v vacancies, and va vacancies arrive to an employed worker. Information about these vacancies is transmitted by the employed worker to a randomly chosen unemployed worker out of his or her h contacts. $(1-u)^h$ is the probability that all the contacts of the employed worker are employed. So with probability $(1-(1-u)^h)$ the employed worker has at least one unemployed contact. Note, that unlike the model by Galeotti and Merlino (2003) the matching through referrals is modeled to be exogenous in the present paper. Job finding rate in case of finding job through referrals is λ_r:

$$\lambda_r = \frac{m_r}{u} \tag{4.9}$$

Job filling rate in the informal market is q_r:

$$q_r = \frac{m_r}{v} \tag{4.10}$$

The unemployment rate is given by the differences between the flows into and out of the unemployment. $1-u$ employed workers lose their job at rate δ, so the flow into the unemployment is $\delta(1-u)$. While u unemployed native workers find a job through formal channels at job finding rate λ_f. So the flow out of the unemployment through the formal channels is $\lambda_f u$. The u unemployed workers may also find a job through referrals at job finding rate λ_r. So the flow out of the unemployment through referrals is $\lambda_r u$. Steady state equations for the unemployment rate can be described by the following equations:

$$\dot{u} = \delta(1-u) - \lambda_f u - \lambda_r u = 0 \tag{4.11}$$

At the steady state the flow into the unemployment equals to the flow out of the unemployment, so the equilibrium condition for unemployment can be written as:

$$u = \frac{\delta}{\delta + \lambda_f + \lambda_r} \tag{4.12}$$

4.3.2 The choice of search intensity

In the model unemployed workers choose the search intensity to maximize the present discounted value of being unemployed. We assume that the there is a cost associated with increasing the search intensity. The cost of s_j units of search φ_j is increasing at the margin and can be expressed by the following formula:

$$\varphi_j(s_j) = C_0 s_j^2 \tag{4.13}$$

Where C_0 is the variable cost of exerting one unit of search effort. We can think of C_0 as of a forgone leisure.

Let w be the wage for all the employed workers, W be the present discounted value of an employed worker, and U be the present discounted value of an unemployed worker. At the exogenous rate δ employed workers lose their job and become unemployed, so the expected capital loss from losing a job for an employed worker is $\delta(W-U)$. Hence the equation for the present discounted value of an employed worker W can be written as:

$$rW = w - \delta(W-U) \quad \text{or} \quad W - U = \frac{w - rU}{(r+\delta)} \tag{4.14}$$

The unemployed workers receive unemployment benefit b while being unemployed. Additionally, supplying s_j units of search intensity costs the worker j, $\varphi_j(s_j)$. At the job finding rate through referrals λ_r the unemployed workers move from into employment through the referral hiring. So the expected capital gain of the unemployed worker from finding a job through referrals is $\lambda_r(W-U)$. While individual unemployed workers move into employment through formal channels at the job finding rate λ_{fj}. Then the expected capital gain of the unemployed worker j from finding a job through formal channels is $\lambda_{fj}(W-U_j)$. Hence the equation for the present discounted value of unemployed worker j will be

$$rU_j = b - C_0 s_j^2 + \lambda_{fj}(W-U_j) + \lambda_r(W-U_j) \tag{4.15}$$

The unemployed worker j chooses the search intensity s_j to maximize U_j taking other variables as given. Thus, the equation for the optimal s_j can be written as

$$-2C_0 s_j + \frac{\partial \lambda_{fj}}{\partial s_j}(W - U_j) = 0 \tag{4.16}$$

Equation (4.16) shows that the expected net gain $(W - U_j)\partial\lambda_{fj}/\partial s_j$ equals to the marginal cost of increasing the search intensity by one unit. In the symmetric equilibrium where the intensity of search of all unemployed workers is the same $s_j = s$, the partial derivative of job finding rate with respect to the search intensity $\partial\lambda_{fj}/\partial s_j$ can be derived using equations (4.4) and (4.6)

$$\frac{\partial \lambda_{fj}}{\partial s_j} = \frac{m_f(su, av)}{su} = \frac{\lambda_f}{s} \tag{4.17}$$

In the symmetric equilibrium the capital gain from finding a job $W - U$ can be derived using equations (4.14) and (4.15).

$$W - U = \frac{w - b + C_0 s^2}{r + \delta + \lambda_f + \lambda_r} \tag{4.18}$$

The derivation of equation (4.18) is presented in Appendix I. After plugging in the expressions for the partial derivative of job finding rate with respect to the search intensity $\partial\lambda_{fj}/\partial s_j$ from (4.17) and the capital gain from finding a job $W - U$ from (4.18) into (4.16), the equation for the optimal search intensity at $s_j = s$ becomes:

$$-2C_0 s + \frac{\lambda_f}{s}\frac{w - b + C_0 s^2}{r + \delta + \lambda_f + \lambda_r} = 0 \tag{4.19}$$

4.3.3 The choice of job advertisement

In the model the firms choose the job advertisement level to maximize the present discounted value of the firms' profit from an open vacancy. We assume that there is a cost associated with increasing the level of job advertisement through formal search channels. The cost of a_i units of job advertisement through formal channels $c(a_i)$:

$$c(a_i) = k + C_0 a_i^2 \tag{4.20}$$

where k is the fixed cost of recruitment through formal channels. Note that the variable cost of the job advertisement equals to the variable cost of exerting one unit of search effort. We can think of k as a fixed cost of placing job advertisements in internet or newspapers, and C_0 as a wage paid to the recruiting staff.

Since workers can be matched to the firm i through two different channels, the expected return of an open vacancy to the firm i consists of two parts. The product of the job filling rate through referrals and the expected net return of the job for the firm i, $q_r(J - V_i)$ is the expected return to the firm i from hiring a worker through referrals. The product of the job filling rate through formal channels and expected net return of the job for the firm

i, $q_{fi}(J - V_i)$ is the expected return to the firm i from hiring a worker through formal channels. Hence the equation for the asset value of an open vacancy for the firm i can be written as:

$$rV_i = -c(a_i) + q_r(J - V_i) + q_{fi}(J - V_i) \tag{4.21}$$

The firm i chooses the level of job advertisement a_i to maximize the asset value of an open vacancy V_i. Thus, the equation for the optimal a_i can be written as

$$-2C_0 a_i + \frac{\partial q_{fi}}{\partial a_i}(J - V_i) = 0 \tag{4.22}$$

where the J is the asset value of a filled job. The equation (22) shows that the expected net profit of the firm i, $(J - V_i)\partial q_{fi}/\partial a_i$ equals to the marginal cost of increasing the job advertisement by one unit. In the symmetric equilibrium where the level of job advertisement of all firms is the same $a_i = a$, the partial derivative of job filling rate with respect to the job advertisement $\partial q_{fi}/\partial a_i$ can be derived using the equations (4.5) and (4.7)

$$\frac{\partial q_{fi}}{\partial a_i} = \frac{m_f(su, av)}{av} = \frac{q_f}{a} \tag{4.23}$$

The equation for the present discounted value of the firm's profit from a filled job:

$$rJ = p - w - \delta(J - V_i) \quad \text{or} \quad J - V_i = \frac{p - w - rV_i}{r + \delta} \tag{4.24}$$

where the firm i benefits from the true productivity of workers p. According to equation (4.24) the cost of a filled job consists of the wage w and the net expected loss from job destruction.

Since the equilibrium is symmetric where the level of job advertisement of all firms is the same $a_i = a$, then $V_i = V$. Thus, under the Free Entry condition the equation for the asset value of an open vacancy can be rewritten as follows:

$$rV = -k - C_0 a^2 + q_r(J - V) + q_f(J - V) = 0 \tag{4.25}$$

And the corresponding equation for the optimal a can be written as:

$$-2C_0 a + \frac{q_f}{a}(J - V) = 0 \tag{4.26}$$

The expression for the asset value of a filled job can be derived using equation (4.25).

$$J = \frac{k + C_0 a^2}{q_r + q_f} \tag{4.27}$$

After plugging in the expressions for the partial derivative of job filling rate with respect to the job advertisement $\partial q_{fi}/\partial a_i$ from (4.23) and the net profit from filling a job $J - V$

from (4.27), the equation for the optimal job advertisement effort becomes:

$$-2C_0a + \frac{q_f}{a}\Big(\frac{k + C_0a^2}{q_r + q_f}\Big) = 0 \tag{4.28}$$

4.3.4 Wage determination

The wage rate is determined using the Nash bargaining rule. An individual firm and worker bargain of the wage taking U and V as given. Which means that an individual firm and worker do not influence the behavior in the rest of the labour market.

$$(W - U)^B (J - V)^{1-B} \longrightarrow \max_w \tag{4.29}$$

where B is the bargaining power of the representative worker. If the negotiation is successful, then the worker is employed and gets the present discounted value of an employed worker W. While if the negotiation is not successful, the worker stays unemployed, so the outside of the worker is the present discounted value of an unemployed worker. Thus, the objective function of the representative worker is the capital gain from finding a job $W - U$. Similar, if the negotiation is successful, and the position is filled, the firm gets the asset value of a filled job J. But if the negotiation is not successful, and the position stays vacant, the firms gets the asset value of an open vacancy. So the objective function of the representative firm is the net return from hiring a worker $J - V$. We can substitute expressions for $W - U$ and $J - U$ from (4.14) and (4.24) respectively and rewrite the Nash bargaining solution.

$$\Big(\frac{w - rU}{r + \delta}\Big)^B \Big(\frac{p - w - rV}{r + \delta}\Big)^{1-B} \longrightarrow \max_w \tag{4.30}$$

When we apply the Free Entry condition $V = 0$ and substitute the expression for rU from (4.14) into the first order condition for the maximization problem we can write the wage equation.

$$w = \frac{Bp(r + \delta + \lambda_f + \lambda_r) + (1 - B)(b - C_0s^2)(r + \delta)}{r + \delta + B\lambda_f + B\lambda_r} \tag{4.31}$$

The derivation of equation for wage (4.31) is presented in Appendix II. We can substitute the RHS of the wage equation (4.31) into the value equation for the present discounted value of the firm's profit from a filled job (4.24) to get the job creation condition:

$$(1 - B)(p - b + C_0s^2) - B\frac{v}{u}(k + C_0a^2) - (r + \delta)\frac{k + C_0a^2}{q_r + q_f} = 0 \tag{4.32}$$

The derivation of job creation condition (4.32) is presented in Appendix III.

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

$$fr = \frac{\lambda_r}{\lambda_f + \lambda_r} \tag{4.33}$$

To sum up, in the model the equilibrium is characterized by the following five conditions:

the equation for the optimal search intensity of unemployed workers, the equation for the optimal job advertisement effort of the firms, the equilibrium condition for unemployment, job creation condition, and fraction of workers who found their job through referrals. The model is used to numerically calculate some values of variables in the model and to find the changes of the key variables over the business cycle.

4.4 Numerical example

Let us first discuss the choice of the values of the exogenous variables and estimated parameters described in Table 4.5. The real productivity of the workers p is normalized to 1. To estimate he matching function in formal market I use the monthly data from the Statistics Department of the Federal Employment Agency from June, 2012 to June, 2017 for 16 German federal states. Estimations show that the matching function demonstrates constant returns to scale. Further, I assume that the variables in the matching function cointegrate, and use Autoregressive distributive lag (ARDL) specification to estimate the parameters of the matching function. While comparing the results obtained using Pooled mean-group (PMG), Mean-group (MG) and fixed effect (FE) models, the Hausman test suggests that FE model is statistically prefered to be used in the estimations (see details in Appendix IV). According to the estimation results the elasticity with respect to the vacancy in the matching function η equals 0.209, and total factor productivity of the matching function λ_0 equals $e^{-0.652} = 0.521$. But since the latter is obtained for the number of total matches, to get the number of matches through the formal search channels only, the total number of matches is multiplied by $1 - fr$.

$$\frac{m_f}{u} = \lambda_f = (1 - fr)e^{-0.652}u^{1-0.209}v^{0.209} = (1 - fr) \cdot 0.521 \cdot u^{0.791}v^{0.209} = \lambda_0 \cdot u^{0.791}v^{0.209} \quad (4.34)$$

Similar to the present study, Stops (2016) and Dengler et al. (2016) estimate the parameters of the matching function for the German labour market. Estimation results of Stops (2016) indicate that the average of the calculated values of λ_0 is around 0.75, while, it is around 1.05 according to the estimation results of Dengler et al. (2016). Using German data Iftikhar and Zaharieva (2019) estimate λ_0 to be 0.5832. So the estimate of this study is close to the lower bound of the values documented in the literature. Whereas, the estimate of the elasticity with respect to the unemployment $(1 - \eta)$ is closer to the upper bound of the findings in the literature. Empirical studies estimated the elasticity with respect to unemployment to be from 0.5 to 0.7. (Petrongolo and Pissarides (2001b)) While, according to Stops (2016) $(1-\eta)$ ranges from 0.570 to 0.797 for different estimation equations. Iftikhar and Zaharieva (2019) estimate $1 - \eta$ to be 0.4379.

The fraction of workers who found job through referrals fr is 0.2306 according to our estimates using IAB data, and 0.3025 when using SOEP data. Alaverdyan (2018) observes the frequency of using referrals to find job for native Germans and immigrants and finds 27.53% of natives and 34.79% of immigrants find their job through referrals. Since we do not distinguish between natives and immigrants, 0.3 is chosen to be close to the average values in the literature and our estimates. Thus, $\lambda_0 = (1 - 0.3) \cdot 0.521 = 0.3647$

The unit period of time being six months, the interest rate r is chosen to be 0.01. The values for the rest of exogenous variables are borrowed from the study by Stupnytska and Zaharieva (2015), since the author chooses the average values in the literature. The values of the exogenous variables is described in Table 4.5.

Variable	Value	Explanation. Source.
p	1	Mean of the workers' true productivity. Normalization.
λ_0	0.3647	Total factor productivity of the matching function. Own calculations
η	0.209	Elasticity with respect to vacancy in the matching function Own calculations.
h	70	Number of the contacts of workers. Stupnytska and Zaharieva (2015).
b	0.5	Unemployment benefit. Average in the literature.
c	0.5	Flow cost of the vacancy. Stupnytska and Zaharieva (2015).
B	0.5	Bargaining power of the workers. Shimer (2005b).
r	0.01	Interest rate. Stupnytska and Zaharieva (2015).
δ	0.2	Job destruction rate. Stupnytska and Zaharieva (2015).

Table 4.5: Values of the exogenous variables

The theoretical model is calibrated with two steps. On the first step the values of exogenous variables, fraction of workers who found job through referrals, vacancy and unemployment rates are plugged into the five equilibrium conditions to solve for α, C_0, k, a and s.The values of these variables are used on the second step to numerically calculate the effect of productivity change on the key parameters of the labour market, such as wages, vacancy and unemployment rates, fraction of workers who found job through referrals, advertisement effort and search effort.

The unemployment rate is calculated according to the ILO guidelines using the SOEP data of the years from 2000 to 2014. The vacancy rate is calculated using the monthly data from the Statistics Department of the Federal Employment Agency from January, 2000 to December, 2013. The ratio of vacancies and unemployment is 0.1048 in the data, so the vacancy rate is $0.1048 \cdot 0.0842 = 0.0088$. Calibration results of the first step are presented in Table 4.6.

Variable	Value	Explanation
fr	0.3	Fraction of workers who found job through referrals
v	0.0088	Vacancy rate
u	0.0842	Unemployment rate
α	0.0616	Vacancy arrival rate
C_0	0.1852	Fixed cost of the advertisement and search effort
k	0.2471	Variable of the recruitment
a	0.2471	Advertisement effort of the firms
s	0.2471	Search effort of the workers

Table 4.6: Calibration results

One the second step the value of the productivity p is changed from 0.55 to 1.5, and the corresponding changes of the labour market parameters are presented in Figure 4.5. According to the Figure 1. over the productivity increase the unemployment rate is decreasing, vacancy rate and the wages are increasing. These findings are inline with the predictions of the Pissarides (2000).

Further, when the productivity increases, the search effort of the workers increases, while the advertisement effort of the firms decreases. When the vacancy rate and wages increase, and the unemployment rate decreases, job search becomes more gainful for the workers. Thus, the increase of the search effort of the workers can be explained by these changes in the labour market. Intuitively, the advertisement of the vacancy should become more gainful for the firms because of the productivity increase. Whereas, the advertisement of the vacancy becomes less gainful for the firms because of the changes, since the wages increase, and there are less unemployed workers available per vacancy issued. The latter negative effect dominates the positive effect of the productivity increase, as a result the advertisement effort decreases.

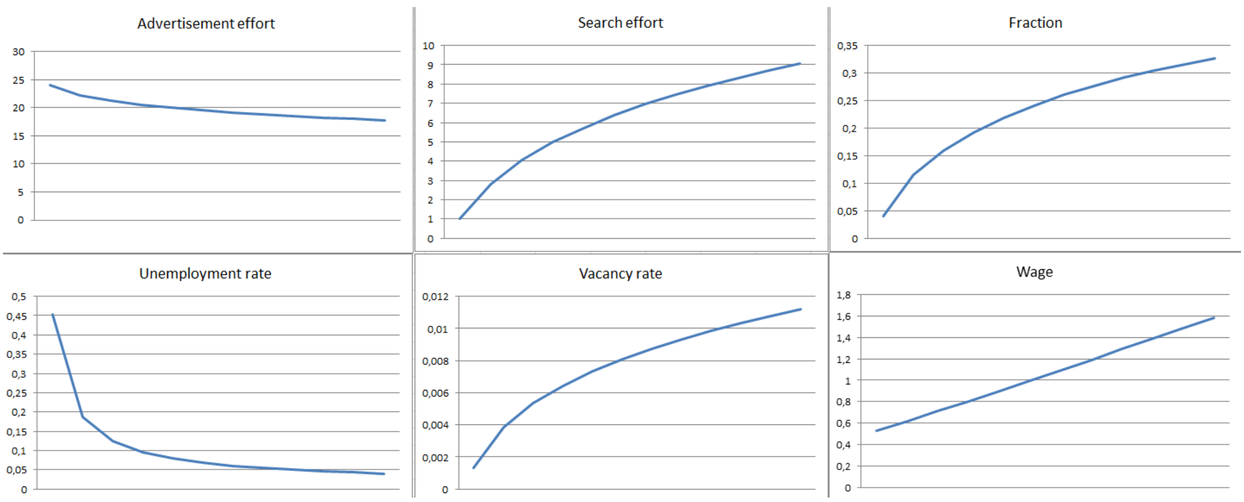

Figure 4.5: Calibration results on the change of the labour market parameters.

On the one hand, as the advertisement effort of the firms in the formal search channels decreases, intuitively, we would expect that the fraction of workers hired through referrals should increase. On the other hand, the search effort of the workers through formal search channels increases, as a result the fraction of workers hired through referrals would decrease. Thus the negative effect of the advertisement effort decrease on the fraction dominates the positive effect of the search effort increase. The calibration results confirm the empirical findings that the firm-side effect on the fraction dominates the effect from the worker's side
.

4.5 Conclusions

Using the two-step estimation procedure we show that in the long run there is a positive correlation between the GDP and the proportion of workers hired through their social networks. Moreover, the results are similar when using both the SOEP and the JVS data for the empirical analysis.

In order to explain the effect of the productivity change on the search and matching channels of firms and workers, this paper presents a search and matching model with two search channels, where the workers choose their search intensity through formal channels, and the firms choose their optimal advertisement intensity.

Calibration results of the model show that during expansions vacancy rate and the wages are increasing, unemployment rate is decreasing. As a result, employment becomes more gainful for the workers, and they increase the job search intensity through formal channels. This reaction of the workers to the labour market conditions reduces the proportion of referral hiring.

When observing the effects of the labour market conditions of the advertisement effort of the firms, on the one hand, the productivity increases, so the advertisement of the vacancy through formal channels should become more gainful for the firms. On the other hand, during expansions there are less unemployed workers available per vacancy issued, and the wages are higher. Thus, the advertisement of the vacancy through formal channels is less gainful for the firms. Since the advertisement effort of the firms through formal channels decreases during expansion, the negative effect on the job advertisement dominates.Which means that according to the behaviour of the firms during expansions the proportion of referral hiring should increase.

As the advertisement effort of the firms in the formal search channels decreases, intuitively, we would expect that the fraction of workers hired through referrals should increase. Whereas, the search effort of the workers through formal search channels increases, as a result the fraction of workers hired through referrals should decrease. The negative effect of the advertisement effort decrease on the fraction or the firm-side effect dominates the positive effect of the search effort increase of the workers or the workers-side effect. The calibration results, that the firm-side effect on the fraction dominates the effect from the worker's side, are inline with the empirical findings.

4.6 Appendix

Appendix I. Derivation of the capital gain of finding a job:

If we substitute the expression for the $W - U$ from the RHS of equation (4.14) into equation (4.15), then the equation for the present discounted value of being unemployed at $s_j = s$ can be written as

$$rU = b - C_0 s^2 + (\lambda_f + \lambda_r)\left(\frac{w - rU}{(r + \delta)}\right) \tag{4.35}$$

When we move the rU from the RHS to the LHS of equation (4.35) we get

$$rU\Big(1 + \frac{\lambda_f + \lambda_r}{(r + \delta)}\Big) = b - C_0 s^2 + \frac{w(\lambda_f + \lambda_r)}{(r + \delta)} \tag{4.36}$$

To get the expression for the rU equation (4.36) can be rewritten as

$$rU = \frac{(b - C_0 s^2)(r + \delta) + w(\lambda_f + \lambda_r)}{r + \delta + \lambda_f + \lambda_r} \tag{4.37}$$

When we substitute the expression for the rU from equation (4.37) into the equation (4.14) we get

$$W - U = \frac{w(r + \delta + \lambda_f + \lambda_r) - (b - C_0 s^2)(r + \delta) - w(\lambda_f + \lambda_r)}{(r + \delta)(r + \delta + \lambda_f + \lambda_r)} \tag{4.38}$$

Equation (4.38) can be simplified to get the following expression for the capital gain of finding a job

$$W - U = \frac{w - b + C_0 s^2}{r + \delta + \lambda_f + \lambda_r} \tag{4.39}$$

Appendix II. Derivation of the Wage equation: From (4.29) we know that the Nash bargaining solution can be expressed as:

$$(W - U)^B (J - V)^{1-B} \longrightarrow \max_w \tag{4.40}$$

The first order condition for the maximization problem we can be written as:

$$\Big(\frac{1}{r + \delta}\Big) B (W - U)^{B-1} (J - V)^{1-B} - \Big(\frac{1}{r + \delta}\Big)(1 - B)(W - U)^B (J - V)^{-B} = 0 \tag{4.41}$$

We can substitute expressions for $W - U$ and $J - U$ from (4.14) and (4.24) respectively and rewrite the first order condition.

$$\Big(\frac{1}{r + \delta}\Big) B \Big(\frac{w - rU}{r + \delta}\Big)^{B-1} \Big(\frac{p - w - rV}{r + \delta}\Big)^{1-B} - \Big(\frac{1}{r + \delta}\Big)(1 - B)\Big(\frac{w - rU}{r + \delta}\Big)^B \Big(\frac{p - w - rV}{r + \delta}\Big)^{-B} = 0 \tag{4.42}$$

If we devide both sides of the equation (4.42) by $(1/r + \delta)((w - rU)/(r + \delta))^{B-1}((p - w - rV)/(r + \delta))^{-B}$

we will get:

$$B(p - w - rV) - (1 - B)(w - rU) = 0 \tag{4.43}$$

From (4.43) the wage w can be expressed as:

$$w = B(p - rV) + (1 - B)rU \tag{4.44}$$

When we apply the Free Entry condition $V = 0$, and substitute the expression for rU from (4.37) into (4.44), we get the following wage equation:

$$w = B(p - rV) + (1 - B)\frac{(b - C_0 s^2)(r + \delta) + w(\lambda_f + \lambda_r)}{r + \delta + \lambda_f + \lambda_r} \tag{4.45}$$

Equation (45) can be simplified to get the following expression for the wage.

$$w = \frac{Bp(r + \delta + \lambda_f + \lambda_r) + (1 - B)(b - C_0 s^2)(r + \delta)}{r + \delta + B\lambda_f + B\lambda_r} \tag{4.46}$$

Appendix III. Derivation of the job creation condition:

In the symmetric equilibrium under the Free Entry condition equation (4.24) can be rewritten as:

$$J = \frac{p - w}{r + \delta} \tag{4.47}$$

Additionally, according to (4.27) the expression for the asset value of a filled job can be written as:

$$J = \frac{k + C_0 a^2}{q_r + q_f} \tag{4.48}$$

When we substitute the expression for the wage from (4.46) into (4.47) we get:

$$\begin{aligned} J = \frac{p(r + \delta + B\lambda_f + B\lambda_r) - Bp(r + \delta + \lambda_f + \lambda_r) - (1 - B)(b - C_0 s^2)(r + \delta)}{(r + \delta + B\lambda_f + B\lambda_r)(r + \delta)} = \\ = \frac{(1 - B)p - (1 - B)(b - C_0 s^2)}{r + \delta + B\lambda_f + B\lambda_r} \end{aligned} \tag{4.49}$$

The RHSs of (4.48) and (4.49) equal each other:

$$\frac{k + C_0 a^2}{q_r + q_f} = \frac{(1 - B)p - (1 - B)(b - C_0 s^2)}{r + \delta + B\lambda_f + B\lambda_r} \tag{4.50}$$

Equation (4.50) can be rewritten as follows:

$$(q_r + q_f)(1 - B)(p - b + C_0 s^2) - B(\lambda_f + \lambda_r)(k + C_0 a^2) - (k + C_0 a^2)(r + \delta) = 0 \tag{4.51}$$

If we divide both sides of (4.51) by $q_r + q_f$ we get the following job creation condition:

$$(1 - B)(p - b + C_0 s^2) - B\frac{v}{u}(k + C_0 a^2) - (r + \delta)\frac{k + C_0 a^2}{q_r + q_f} = 0 \tag{4.52}$$

where

$$\frac{\lambda_f + \lambda_r}{q_r + q_f} = \frac{\frac{m_f(su,av)}{u} + \frac{m_r}{u}}{\frac{m_f(su,av)}{v} + \frac{m_r}{v}} = \frac{v}{u} \tag{4.53}$$

Appendix IV. Estimation of the matching function in formal market:

We follow the random matching approach and assume that matching function in formal market has the following Cobb-Douglass form:

$$m_f(su, av) = \lambda_0 (su)^{\alpha} (av)^{\beta} \tag{4.54}$$

where α is the elasticity with respect to the unemployment in the matching function, and β elasticity with respect to the vacancy rate in the matching function. We can linearize the equation by taking the logarithms.

$$\log(m_f(su, av)) = \log \lambda_0 + \alpha \log(su) + \beta \log(av) \tag{4.55}$$

The parameters λ_0, α and β can be estimated using the following estimation equation.

$$\log m_{i,t} = C_0 + B_u \log u_{i,t} + B_v \log v_{i,t} + \epsilon_{i,t} \tag{4.56}$$

where $C_0 = \log \lambda_0$, $B_u = \alpha$ and $B_v = \beta$. $m_{i,t}$ is the flow from unemployment to employment in the state i at time period t, $u_{i,t}$ and $v_{i,t}$ are number of unemployed workers and vacancies in the state i at time period t respectively, and $\epsilon_{i,t}$ is the error term for the state i and time period t. In the estimations we use the monthly data from the Statistics Department of the Federal Employment Agency from June, 2012 to June, 2017 for 16 German federal states. The results of OLS estimation of equation (4.56) are presented in Table 4.7.

When we include time trend t in the estimation equation we get the following equation.

$$\log m_{i,t} = C_0 + B_u \log u_{i,t} + B_v \log v_{i,t} + t + \epsilon_{i,t} \tag{4.57}$$

The t test suggests that the time trend should not be added to the regression equation (4.56). Next, we control for the state fixed effects in the equation (4.3). The equation can be rewritten as following:

$$\log m_{i,t} = C_i + B_u \log u_{i,t} + B_v \log v_{i,t} + \epsilon_{i,t} \tag{4.58}$$

where C_i is the state specific intercept. Equation (4.58) with a time trend can be rewritten as following:

$$\log m_{i,t} = C_i + B_u \log u_{i,t} + B_v \log v_{i,t} + t + \epsilon_{i,t} \tag{4.59}$$

The t test suggests that the time trend should be added to the regression equation (4.58). The estimation results of the four equations mentioned above are presented in Table 4.7.

Further, Hadri test for unit roots shows that the null hypothesis that all panels are stationary is rejected, so at least some panels have unit root for $\log m_{i,t}$, $\log u_{i,t}$ and $\log v_{i,t}$.

Table 4.7: Estimation results of OLS and FE regressions.

Regression equation	(4.56)	(4.57)	(4.58)	(4.59)
Estimation model	OLS	OLS	FE	FE
Dependent variable	$logm_{i,t}$	$logm_{i,t}$	$logm_{i,t}$	$logm_{i,t}$
$logu_{i,t}$	0.778***	0.776***	0.610***	0.599***
	(0.0145)	(0.0165)	(0.0993)	(0.0988)
$logv_{i,t}$	0.234***	0.236***	0.132**	-0.0436
	(0.0123)	(0.0141)	(0.0407)	(0.0657)
t		-0.0000948		0.00177***
		(0.000325)		(0.000524)
Constant	-1.208***	-1.139***	1.794	2.520
	(0.0844)	(0.2508)	(1.519)	(1.525)
Observations	976	976	976	976
R^2	0.9604	0.9604	0.9596	0.9398

Standard errors in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

But the same test shows that the stationarity of the first time differences of the variables can not be rejected. Thus, the time series of the variables are integrated of order one and they are likely to be cointegrated. So, if I assume that the variables cointegrate, Pooled mean-group (PMG) model by Pesaran et al. (1999) can be applied.

The long-run matching function is given by:

$$\log m_{i,t} = C_i + B_{ui} \log u_{i,t} + B_{vi} \log v_{i,t} + v_{i,t} \tag{4.60}$$

Autoregressive distributive lag (ARDL) specification corresponding to equation (4.60):

$$\log m_{i,t} = \lambda_i \log m_{i,t-1} + C_i + \delta_{10i} \log u_{i,t} + \delta_{20i} \log v_{i,t} + \delta_{11i} \log u_{i,t-1} + \delta_{21i} \log v_{i,t-1} + \epsilon_{i,t} \tag{4.61}$$

If we add $-\log m_{i,t-1}$ on both sides of equation (4.61) and rearrange it we can get:

$$\begin{aligned}\Delta \log m_{i,t} = -(1-\lambda_i)(\log m_{i,t-1} - \frac{\delta_{10i}+\delta_{11i}}{1-\lambda_i} \log u_{i,t} - \frac{\delta_{20i}+\delta_{21i}}{1-\lambda_i} \log v_{i,t}) \\ - \delta_{11i} \Delta \log u_{i,t} - \delta_{21i} \Delta \log v_{i,t} + C_i + \epsilon_{i,t}\end{aligned} \tag{4.62}$$

or

$$\begin{aligned}\Delta \log m_{i,t} = \phi_i(\log m_{i,t-1} - B_{ui} \log u_{i,t} - B_{vi} \log v_{i,t}) \\ - \delta_{11i} \Delta \log u_{i,t} - \delta_{21i} \Delta \log v_{i,t} + C_i + \epsilon_{i,t}\end{aligned} \tag{4.63}$$

Where the error-correcting speed of adjustment term $\phi_i = -(1-\lambda_i)$, and long-run relationships $B_{ui} = \frac{\delta_{10i}+\delta_{11i}}{1-\lambda_i}$, $B_{vi} = \frac{\delta_{20i}+\delta_{21i}}{1-\lambda_i}$. There is a long-run relationship between the variables when ϕ_i is significantly negative, meaning that variables show a return to a long-run equilibrium. While there is no long-run relationship between variables when $\phi_i = 0$.

In the PMG model the long-run coefficients are constrained to be equal across groups, i.e. $B_{ui} = B_u$ and $B_{vi} = B_v$, but the short run coefficients and the error variances are allowed to differ across groups. Thus, during estimation of the PMG model we use the following parametrization.

$$\begin{aligned}\Delta \log m_{i,t} = \phi_i(\log m_{i,t-1} - B_u \log u_{i,t} - B_v \log v_{i,t}) \\ - \delta_{11i}\Delta \log u_{i,t} - \delta_{21i}\Delta \log v_{i,t} + C_i + \epsilon_{i,t}\end{aligned} \tag{4.64}$$

In the mean-group (MG) model developed by Pesaran and Smith (1995) the model is fitted for each group separately, so all the coefficients are allowed to vary across groups and the means of them are reported. The parametrization of the MG is the following.

$$\begin{aligned}\Delta \log m_{i,t} = \phi_i(\log m_{i,t-1} - B_{ui} \log u_{i,t} - B_{vi} \log v_{i,t}) \\ - \delta_{11i}\Delta \log u_{i,t} - \delta_{21i}\Delta \log v_{i,t} + C_i + \epsilon_{i,t}\end{aligned} \tag{4.65}$$

Fixed-effects (FE) estimation approach constrains all the slope coefficients to be equal across groups, i.e. $B_{ui} = B_u$, $B_{vi} = B_v$, $\delta_{11i} = \delta_{11}$, and $\delta_{21i} = \delta_{21}$, and only the intercept and the error variances are allowed be different across groups.

$$\begin{aligned}\Delta \log m_{i,t} = \phi(\log m_{i,t-1} - B_u \log u_{i,t} - B_v \log v_{i,t}) \\ - \delta_{11}\Delta \log u_{i,t} - \delta_{21}\Delta \log v_{i,t} + C_i + \epsilon_{i,t}\end{aligned} \tag{4.66}$$

The estimation results of the PMG, MG, and FE models are presented in Table 4.8.

The results in Table 4.8 show that the error-correcting speed of adjustment term ϕ_i is negative and statistically significant, so the variables show a return to a long-run equilibrium. Also, for all three models the long-run coefficient of time variable is statistically insignificant. Thus, Hausman test was conducted to determine which of the models without time trend is statistically preferred, and the test shows that FE model is more preferable. Furthermore, the test of H_0: $B_u + B_v = 1$ did not reject the null hypothesis of the constant return to scale in the FE model without time trend. So we can rewrite equation (4.54) so that it has constant returns to scale (CRS):

$$m_f(su, av) = \lambda_0 (su)^{1-\eta}(av)^{\eta} \tag{4.67}$$

If we divide both sides of (67) by su we get:

$$\frac{m_f(su, av)}{su} = \lambda_0 \frac{(av)^{\eta}}{(su)^{\eta}} \tag{4.68}$$

Table 4.8: Estimation results of the PMG, MG, and FE models.

Regression equation	(4.64)		(4.65)		(4.66)	
Estimation model	PMG	PMG	MG	MG	FE	FE
Dependent variable	$\Delta \log m_{i,t}$	$\Delta \log m_{i,t}$	$\Delta \log m_{i,t}$	$\Delta \log m_{i,t}$	$\Delta \log m_{i,t}$	$\Delta \log m_{i,t}$
Long-run coef.						
B_u	0.730***	0.727***	0.372*	0.367*	0.788***	0.786***
	(0.172)	(0.171)	(0.172)	(0.154)	(0.0895)	(0.0884)
B_v	0.168**	0.136	-0.0159	-0.00807	0.145**	0.110
	(0.0589)	(0.0842)	(0.0826)	(0.0797)	(0.0441)	(0.0578)
t		0.000329		0.000244		0.000368
		(0.000631)		(0.000729)		(0.000453)
Short-run coef.						
ϕ	-0.616***	-0.618***	-0.642***	-0.643***	-0.580***	-0.581***
	(0.0297)	(0.0298)	(0.0329)	(0.0329)	(0.0320)	(0.0321)
δ_{11}	-2.953***	-2.922***	-2.828***	-2.817***	-2.890***	-2.854***
	(0.227)	(0.224)	(0.245)	(0.262)	(0.251)	(0.265)
δ_{21}	-1.318***	-1.332***	-1.280***	-1.327***	-1.080***	-1.088***
	(0.118)	(0.119)	(0.148)	(0.168)	(0.141)	(0.145)
Constant	0.0149	0.097***	3.478	3.649*	-0.247	-0.1717
	(0.0229)	(0.0273)	(1.908)	(1.615)	(0.829)	(0.829)
Observations	960	960	960	960	960	960

Standard errors in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Equation (68) can be linearized by taking the logarithms.

$$\log mu = \log \lambda_0 + \eta \log vu \tag{4.69}$$

where $\log mu = \log \frac{m_f(su,av)}{su}$ and $\log vu = \log \frac{(av)^\eta}{(su)^\eta}$.

The parameters $\log \lambda_0 = C_0$ and $\eta = B$ can be estimated using the following estimation equation.

$$\log mu_{i,t} = C_0 + B \log u_{i,t} + \epsilon_{i,t} \tag{4.70}$$

When we include time trend t in the estimation equation we get the following equation.

$$\log mu_{i,t} = C_0 + B \log u_{i,t} + t + \epsilon_{i,t} \tag{4.71}$$

If we control for the state fixed effects in the equation (71), the equation can be rewritten as following:

$$\log mu_{i,t} = C_i + B \log u_{i,t} + \epsilon_{i,t} \tag{4.72}$$

where C_i is again the state specific intercept. Equation (72) with a time trend can be rewritten as following:

$$\log mu_{i,t} = C_i + B \log u_{i,t} + t + \epsilon_{i,t} \tag{4.73}$$

The estimation results of the four equations mentioned above are presented in Table 4.9. Next, we conduct a Hadri test for unit root and find that the null hypothesis that all panels

Table 4.9: Estimation results of OLS and FE regressions with CRS matching function.

Regression equation	(4.70)	(4.71)	(4.72)	(4.73)
Estimation model	OLS	OLS	FE	FE
Dependent variable	$logmu_{i,t}$	$logmu_{i,t}$	$logmu_{i,t}$	$logmu_{i,t}$
$logvu_{i,t}$	0.236***	0.239***	0.202***	0.121**
	(0.0122)	(0.0140)	(0.0180)	(0.0386)
t		-0.000165		0.00116*
		(0.000322)		(0.000486)
Constant	-1.066***	-0.951***	-1.125***	-2.030***
	(0.0217)	(0.2251)	(0.0314)	(0.3808)
Observations	976	976	976	976
R^2	0.2769	0.2771	0.2769	0.2495

Standard errors in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

are stationary is rejected, so at least some panels have unit root for $\log mu_{i,t}$, and $\log vu_{i,t}$. But the same test shows that the stationarity of the first time differences of the variables can not be rejected. Thus, the time series of the variables are integrated of order one and they are likely to be cointegrated. If I assume that the variables cointegrate, Pooled mean-group (PMG) model can be applied. The long-run matching function is given by:

$$\log mu_{i,t} = C_i + B \log vu_{i,t} + \upsilon_{i,t} \tag{4.74}$$

Autoregressive distributive lag (ARDL) specification corresponding to the equation (4.74):

$$\log mu_{i,t} = \lambda_i \log mu_{i,t-1} + C_i + \delta_{1i} \log vu_{i,t} + \delta_{2i} \log vu_{i,t-1} + \epsilon_{i,t} \tag{4.75}$$

If we add $-\log mu_{i,t-1}$ on both sides of equation (4.75) and rearrange it we can get:

$$\Delta \log mu_{i,t} = -(1-\lambda_i)(\log mu_{i,t-1} - \frac{\delta_{1i} + \delta_{2i}}{1-\lambda_i} \log vu_{i,t}) - \delta_{2i} \Delta \log vu_{i,t} + C_i + \epsilon_{i,t} \tag{4.76}$$

or

$$\Delta \log mu_{i,t} \quad = \quad \phi_i(\log mu_{i,t-1} \ - \ B_i \log vu_{i,t}) \ - \ \delta_{2i} \Delta \log vu_{i,t} \ + \ C_i \ + \ \epsilon_{i,t} \tag{4.77}$$

Where $\phi_i = -(1-\lambda_i)$, $B_i = \frac{\delta_{1i}+\delta_{2i}}{1-\lambda_i}$.

In the PMG model the long-run coefficient is constrained to be equal across groups $B_i = B$, but the short run coefficient and the error variances are allowed to differ across groups. Thus, during estimation of the PMG model we use the following parametrization.

$$\Delta \log mu_{i,t} \quad = \quad \phi_i(\log mu_{i,t-1} \ - \ B \log vu_{i,t}) \ - \ \delta_{2i} \Delta \log v_{i,t} \ + \ C_i \ + \ \epsilon_{i,t} \tag{4.78}$$

In the mean-group (MG) model the model is fitted for each group separately, so all the coefficients are allowed to vary across groups. The parametrization of the MG is the following.

$$\Delta \log mu_{i,t} \quad = \quad \phi_i(\log mu_{i,t-1} \ - \ B_i \log vu_{i,t}) \ - \ \delta_{2i} \Delta \log v_{i,t} \ + \ C_i \ + \ \epsilon_{i,t} \tag{4.79}$$

Fixed-effects (FE) estimation approach constrains all the slope coefficients to be equal across groups, i.e. $B_i = B$, and $\delta_{2i} = \delta_2$, and only the intercept and the error variances are allowed be different across groups.

$$\Delta \log mu_{i,t} \quad = \quad \phi(\log mu_{i,t-1} \ - \ B \log vu_{i,t}) \ - \ \delta_2 \Delta \log v_{i,t} \ + \ C_i \ + \ \epsilon_{i,t} \tag{4.80}$$

The estimation results of the PMG, MG, and FE models with CRS matching function are presented in Table 4.10.

The results in Table 4.10 show that the error-correcting speed of adjustment term ϕ_i is negative and statistically significant. This means that the variables show a return to a long-run equilibrium. Moreover, for all three models the long-run coefficient of time variable is statistically insignificant. To compare the models we used the models without time trend. Hausman test was conducted to determine which of the models is statistically

Table 4.10: Estimation results of the PMG, MG, and FE models with CRS matching function.

Regression equation	(4.78)		(4.79)		(4.80)	
Estimation model	PMG	PMG	MG	MG	FE	FE
Dependent variable	$\Delta \log mu_{i,t}$	$\Delta \log mu_{i,t}$	$\Delta \log mu_{i,t}$	$\Delta \log mu_{i,t}$	$\Delta \log mu_{i,t}$	$\Delta \log mu_{i,t}$
Long-run coef.						
B	0.230***	0.202***	0.241***	0.228***	0.209***	0.165***
	(0.0291)	(0.0530)	(0.0322)	(0.0467)	(0.0181)	(0.0302)
t		0.000391		0.000338		0.000643
		(0.000604)		(0.000813)		(0.000549)
Short-run coef.						
ϕ	-0.626***	-0.627***	-0.640***	-0.644***	-0.591***	-0.594***
	(0.0315)	(0.0313)	(0.0325)	(0.0324)	(0.0313)	(0.0312)
δ_{2i}	-0.565***	-0.588***	-0.564***	-0.608***	-0.518***	-0.551***
	(0.0852)	(0.0826)	(0.0984)	(0.120)	(0.0768)	(0.0987)
Constant	-0.666***	-0.859***	-0.667***	-0.862*	-0.652***	-0.951***
	(0.0378)	(0.0460)	(0.0534)	(0.3534)	(0.0351)	(0.2373)
Observations	960	960	960	960	960	960

Standard errors in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

preferred, and the test shows that FE model is more preferable. So, we use the coefficients of the FE model without time trend and with CRS matching function (4.79) to find the parameters needed for numerical calculations.

Bibliography

Addison, J. T. and P. Portugal (2002). Job search methods and outcomes. *Oxford Economic Papers 54*(3), 505–533.

Alaverdyan, S. (2018). Why do migrant workers rely more often on referrals? *Bielefeld Working Papers in Economics and Management 05*.

Aleksynska, M. and A. Tritah (2013). Occupation-education mismatch of immigrant workers in europe: Context and policies. *Economics of Education Review 36*, 229–244.

Allen, J. and E. De Weert (2007). What do educational mismatches tell us about skill mismatches? a cross-country analysis. *European Journal of Education 42*(1), 59–73.

Andersen, E. B. (1970). Asymptotic properties of conditional maximum-likelihood estimators. *Journal of the Royal Statistical Society: Series B (Methodological) 32*(2), 283–301.

Andersen, E. B. (1971). The asymptotic distribution of conditional likelihood ratio tests. *Journal of the American Statistical Association 66*(335), 630–633.

Battu, H., P. Seaman, and Y. Zenou (2011). Job contact networks and the ethnic minorities. *Labour Economics 18*(1), 48–56.

Behtoui, A. (2008). Informal recruitment methods and disadvantages of immigrants in the swedish labour market. *Journal of Ethnic and Migration Studies 34*(3), 411–430.

Bentolila, S., C. Michelacci, and J. Suarez (2010). Social contacts and occupational choice. *Economica 707*(305), 20–45.

Brown, M., E. Setren, and G. Topa (2016). Do informal referrals lead to better matches? evidence from a firm's employee referral system. *Journal of Labor Economics 34*(1), 161–209.

Butler, J. S. and R. Moffitt (1982). A computationally efficient quadrature procedure for the one-factor multinomial probit model. *Econometrica: Journal of the Econometric Society*, 761–764.

Cahuc, P., S. Carcillo, and A. Zylberberg (2014). *Labor economics*. MIT press.

Cahuc, P. and F. Fontaine (2009). On the efficiency of job search with social networks. *Journal of Public Economic Theory 11*(3), 411–439.

Calvo-Armengol, A. and M. O. Jackson (2004). The effects of social networks on employment and inequality. *The American Economic Review 94*(3), 426–454.

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

Cappellari, L. and K. Tatsiramos (2015). With a little help from my friends? quality of social networks, job finding and job match quality. *European Economic Review 78*, 55–75.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *The Review of Economic Studies 47*(1), 225–238.

Chiswick, B. R. and P. W. Miller (2008). Why is the payoff to schooling smaller for immigrants? *Labour Economics 15*(6), 1317–1340.

Chiswick, B. R. and P. W. Miller (2010). The effects of educational-occupational mismatch on immigrant earnings in australia, with international comparisons. *International Migration Review 44*(4), 869–898.

Cingano, F. and A. Rosolia (2012). People i know: job search and social networks. *Journal of Labor Economics 30*(2), 291–332.

Dawid, H. and S. Gemkow (2014). How do social networks contribute to wage inequality? insights from an agent-based analysis. *Industrial and Corporate Change 23*(5), 1171–1200.

DeGroot, M. H. (2005). *Optimal statistical decisions*, Volume 82. John Wiley & Sons.

Dengler, K., M. Stops, and B. Vicari (2016). Occupation-specific matching efficiency. Technical report, IAB-Discussion Paper.

Diamond, P. A. (1982). Wage determination and efficiency in search equilibrium. *The Review of Economic Studies 49*(2), 217–227.

Drever, A. I. and O. Hoffmeister (2008). Immigrants and social networks in a job-scarce environment: The case of germany. *International Migration Review 42*(2), 425–448.

Dustmann, C., A. Glitz, U. Schönberg, and H. Brücker (2016). Referral-based job search networks. *The Review of Economic Studies 83*(2), 514–546.

Eisnecker, P. and D. Schacht (2016). Half of the refugees in germany found their first job through social contacts. *DIW Economic Bulletin 6*(34/35), 414–421.

Elliot, J. (2001). Referral hiring and ethnically homogeneous jobs: How prevalent is the connection and for whom? *Social Science Research 30*, 401–425.

Galenianos, M. (2013). Learning about match quality and the use of referrals. *Review of Economic Dynamics 16*(4), 668–690.

Galeotti, A. and L. P. Merlino (2003). Endogenous job contact networks. *International Economic Review 55*(4), 1201–1226.

Ganzeboom, H. B., P. M. De Graaf, and D. J. Treiman (1992). A standard international socio-economic index of occupational status. *Social Science Research 21*(1), 1–56.

Glitz, A. (2017). Coworker networks in the labour market. *Labour Economics 44*, 218–230.

Granovetter, M. (1995). Getting a job (revised edition). *Chicago: University of Chicago Press 1*, 995.

Granovetter, M. S. (1974). *Getting a Job : A study of Contacts and Careers.* Cambridge, Mass,.

Greene, W. H. (2007). *Econometric analysis. 6th. ed.* Upper Saddle River, NJ: Pearson/Prentice Hall.

Hagedorn, M. and I. Manovskii (2008). The cyclical behavior of equilibrium unemployment and vacancies revisited. *American Economic Review 98*(4), 1692–1706.

Hall, R. E. and P. R. Milgrom (2008). The limited influence of unemployment on the wage bargain. *American economic review 98*(4), 1653–74.

Helberger, C. (1988). Eine überprüfung der linearitätsannahme der humankapitaltheorie. *Bildung, Beruf, Arbeitsmarkt 151*, 170.

Holzer, H. J. (1988). Search method use by unemployed youth. *Journal of labor economics 6*(1), 1–20.

Horvath, G. (2014). Occupational mismatch and social networks. *Journal of Economic Behavior and Organization 106*(3), 442–468.

Iftikhar, Z. and A. Zaharieva (2019). General equilibrium effects of immigration in germany: Search and matching approach. *Review of Economic Dynamics 31*, 245–276.

Ioannides, Y. M. and L. Datcher Loury (2004). Job information networks, neighborhood effects, and inequality. *Journal of economic literature 42*(4), 1056–1093.

Jackson, M. O. (2010). *Social and economic networks.* Princeton university press.

Jann, B. et al. (2008). A stata implementation of the blinder-oaxaca decomposition. *Stata Journal 8*(4), 453–479.

Johnson, N. L. and S. Kotz (1972). Distributions in statiscs; continuous multivariate distributions. Technical report.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of political economy 87*(5, Part 1), 972–990.

Jovanovic, B. (1984). Matching, turnover, and unemployment. *Journal of political Economy 92*(1), 108–122.

Kaas, L. and C. Manger (2012). Ethnic discrimination in germany's labour market: a field experiment. *German Economic Review 13*(1), 1–20.

Kugler, A. D. (2003). Employee referrals and efficiency wages. *Labour Economics 10*, 531–556.

Lancaster, T. (2000). The incidental parameter problem since 1948. *Journal of econometrics 95*(2), 391–413.

Lancee, B. (2016). Job search methods and immigrant earnings: A longitudinal analysis of the role of bridging social capital. *Ethnicities 16*(3), 349–367.

Maddala, G. S. (1987). Limited dependent variable models using panel data. *Journal of Human resources*, 307–338.

Marsden, P. V. and E. H. Gorman (2001). Social networks, job changes and recruitment. *Sourcebook of Labour Markets: Evolving Structures and Processes 19*, 467–502.

McPherson, M., L. Smith-Lovin, and J. M. Cook (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology 27*(1), 415–444.

Montgomery, J. D. (1991). Social networks and labor-market outcomes: Toward an economic analysis. *The American Economic Review 81*(5), 1408–1418.

Mortensen, D. T. and C. A. Pissarides (1994). Job creation and job destruction in the theory of unemployment. *Review of Economic Studies 61*, 397–415.

Mortensen, D. T. and C. A. Pissarides (1999). Job reallocation, employment fluctuations and unemployment. *Handbook of macroeconomics 1*, 1171–1228.

Moulton, B. R. (1986). Random group effects and the precision of regression estimates. *Journal of econometrics 32*(3), 385–397.

Naylor, J. C. and A. F. Smith (1982). Applications of a method for the efficient computation of posterior distributions. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 31*(3), 214–225.

Neyman, J., E. L. Scott, et al. (1948). Consistent estimates based on partially consistent observations. *Econometrica 16*(1), 1–32.

Ooka, E. and B. Wellman (2006). Does social capital pay off more within or between ethnic groups? analyzing job searchesin five toronto ethnic groups. *Inside the Mosaic Chapter 9*, University of Toronto Press.

Pellizzari, M. (2010). Do friends and relatives really help in getting a good job? *Industrial and Labor Relations Review 63*(3), 494–510.

Pesaran, M. H., Y. Shin, and R. P. Smith (1999). Pooled mean group estimation of dynamic heterogeneous panels. *Journal of the American Statistical Association 94*(446), 621–634.

Pesaran, M. H. and R. Smith (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of econometrics 68*(1), 79–113.

Petrongolo, B. and C. A. Pissarides (2001a). Looking into the black box: A survey of the matching function. *Journal of Economic literature 39*(2), 390–431.

Petrongolo, B. and C. A. Pissarides (2001b). Looking into the black box: A survey of the matching function. *Journal of Economic literature 39*(2), 390–431.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The american economic review 62*(4), 659–661.

Piracha, M. and F. Vadean (2013). Migrant educational mismatch and the labour market. *International Handbook on the Economics of Migration*, 176–192.

Pissarides, C. A. (2000). *Equilibrium unemployment theory* (2 ed.). Cambridge, Mass.: MIT Press.

Pistaferri, L. (1999). Informal networks in the italian labor market. *Giornale degli Economisti e Annali di Economia 58*(3-4), 355–375.

Rasch, G. (1960). Studies in mathematical psychology: I. probabilistic models for some intelligence and attainment tests.

Rebien, M., M. Stops, and A. Zaharieva (2017). Formal search and referrals from a firm's perspective. *IMW working paper 517*.

Robst, J. (2007). Education, college major, and job match: Gender differences in reasons for mismatch. *Education Economics 15*(2), 159–175.

Schwarze, J. et al. (1991). Ausbildung und einkommen von männern: Einkommensfunktionsschätzungen für die ehemalige ddr und die bundesrepublik deutschland. *Mitteilungen aus der Arbeitsmarkt-und Berufsforschung 24*(1), 63–69.

Shimer, R. (2004). Search intensity. Technical report, mimeo, University of Chicago.

Shimer, R. (2005a). The cyclical behavior of equilibrium unemployment and vacancies. *American economic review 95*(1), 25–49.

Shimer, R. (2005b). The cyclical behavior of equilibrium unemployment and vacancies. *American economic review 95*(1), 25–49.

Simon, C. J. and J. T. Warner (1992). Matchmaker, matchmaker: The effect of old boy networks on job match quality, earnings, and tenure. *Journal of Labor Economics 10*(3), 306–330.

Solon, G., R. Barsky, and J. A. Parker (1994). Measuring the cyclicality of real wages: how important is composition bias? *The quarterly journal of economics 109*(1), 1–25.

StataCorp, L. (2013). Stata 13 base reference manual. *College Station, Texas: Stata*.

Stops, M. (2016). Revisiting german labour market reform effects—a panel data analysis for occupational labour markets. *IZA Journal of European Labor Studies 5*(1), 14.

Stupnytska, Y. (2015). Asymmetric information in a search model with social contacts.

Stupnytska, Y. and A. Zaharieva (2015). Explaining u-shape of the referral hiring pattern in a search model with heterogeneous workers. *Journal of Economic Behavior and Orgainization 119*, 211–233.

Stupnytska, Y. and A. Zaharieva (2017). Optimal policy and the role of social contacts in a search model with heterogeneous workers. *Journal of Public Economic Theory 19*(5), 957–985.

Sylos Labini, M. (2005). Social networks and wages: It is all about connections! Technical report, LEM Working Paper Series.

Titzmann, P. F. and R. K. Silbereisen (2009). Friendship homophily among ethnic german immigrants: A longitudinal comparison between recent and more experienced immigrant adolescents. *Journal of family psychology 23*(3), 301.

Tumen, S. (2016). Informal versus formal search: Which yields better pay? *International Journal of Economic Theory 12*(3), 257–277.

Wolbers, M. H. (2003). Job mismatches and their labour-market effects among school-leavers in europe. *European Sociological Review 19*(3), 249–266.

Zaharieva, A. (2018). On the optimal diversification of social networks in frictional labour markets with occupational mismatch. *Labour Economics 50*, 112–127.