Universität Bielefeld

Faculty of Business Administration and Economics

Working Papers in Economics and Management

No. 4-2019
March 2019

Immigration, Social Networks and Occupational Mismatch

Sevak Alaverdyan, Anna Zaharieva

Bielefeld University
P.O. Box 10 01 31
33501 Bielefeld – Germany

ISSN 2196–2723

ABCD www.wiwi.uni–bielefeld.de

Immigration, Social Networks and Occupational Mismatch

Sevak Alaverdyan [*], Anna Zaharieva[†]

March 26, 2019

Abstract

In this study we investigate the link between the job search channels that workers use to find employment and the probability of occupational mismatch in the new job. Our specific focus is on differences between native and immigrant workers. We use data from the German Socio-Economic Panel (SOEP) over the period 2000-2014. First, we document that referral hiring via social networks is the most frequent single channel of generating jobs in Germany; in relative terms referrals are used more frequently by immigrant workers compared to natives. Second, our data reveals that referral hiring is associated with the highest rate of occupational mismatch among all channels in Germany. 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 occupations. When modeling social networks we take into account ethnic and professional homophily in the link formation. Our model predicts that immigrant workers face stronger risk of unemployment and often rely on recommendations from their friends and relatives as a channel of last resort. Furthermore, higher rates of referral hiring produce more frequent occupational mismatch of the immigrant population compared to natives. We test this prediction empirically and confirm that more intensive network hiring contributes significantly to higher rates of occupational mismatch among immigrants. Finally, we document that the gaps in the incidence of referrals and mismatch rates are reduced among second generation immigrants indicating some degree of integration in the German labour market.

Keywords: job search, referrals, social networks, occupational mismatch, immigration

JEL Classification: J23, J31, J38, J64

[*]Bielefeld Graduate School of Economics and Management, Bielefeld University, 33501 Bielefeld, Germany.

[†]Faculty of Business Administration and Economics, Center for Mathematical Economics, Bielefeld University, 33501 Bielefeld, Germany. Email: azaharieva@wiwi.uni-bielefeld.de

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[1], 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

[1] Wolbers (2003), Allen and De Weert (2007), Robst (2007)

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.

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%.

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.

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. 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.

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 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 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 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.[2] 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 2. The detailed estimation results with the coefficients of all variables are presented in Table 13 in Appendix I.

Table 2: Employment rates: logistic regression

Variables	Dependent variable: EMP							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EDU	0.315***	0.314***	0.315***	0.324***	0.371***	0.371***	0.368***	0.344***
	(76.14)	(76.09)	(76.17)	(76.99)	(80.91)	(80.81)	(80.55)	(75.31)
AGE		0.00225**	0.00225**	-0.0110***	-0.00943***	-0.00971***	-0.0171***	-0.0186***
		(3.24)	(3.25)	(-13.46)	(-11.28)	(-11.56)	(-18.52)	(-20.06)
FEMALE			-0.138***	-0.107***	-0.113***	-0.115***	-0.122***	-0.134***
			(-8.71)	(-6.59)	(-6.86)	(-6.98)	(-7.37)	(-8.09)

Continued on next page

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

Variables				Dependent variable: EMP				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MARST(Reference: Married)								
[2] Single				-0.867***	-0.769***	-0.774***	-0.957***	-1.003***
				(-39.95)	(-34.43)	(-34.54)	(-39.67)	(-41.38)
[3] Widowed				-0.323***	-0.300***	-0.298***	-0.346***	-0.370***
				(-5.60)	(-5.15)	(-5.12)	(-5.93)	(-6.33)
[4] Divorced				-0.770***	-0.743***	-0.747***	-0.807***	-0.839***
				(-32.41)	(-30.85)	(-30.94)	(-33.04)	(-34.18)
[5] Separated				-0.752***	-0.748***	-0.748***	-0.801***	-0.815***
				(-16.98)	(-16.62)	(-16.60)	(-17.70)	(-17.94)
NCHILD							-0.175***	-0.167***
							(-20.25)	(-19.22)
MIG								-0.652***
								(-24.68)
STATE					v	v	v	v
Survey year t						v	v	v
Constant	-1.311***	-1.406***	-1.349***	-0.542***	-1.115***	-1.050***	-0.520***	-0.125
	(-28.11)	(-25.53)	(-24.25)	(-9.04)	(-14.31)	(-12.72)	(-6.01)	(-1.42)
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
Standard errors are in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$								

Table 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. 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[3]. 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 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.

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

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

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: 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, 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 4. While the detailed estimation results with the coefficients of all variables are presented in Table 14 in Appendix II.

Table 4: Estimation results of referral hiring.

Variables	Dependent variable: REF									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
EDU	-0.106***	-0.106***	-0.0988***	-0.0798***	-0.0707***	-0.0641***	-0.0643***	-0.0649***	-0.0291***	-0.0290***
	(-17.48)	(-17.33)	(-16.05)	(-12.68)	(-10.80)	(-9.67)	(-9.62)	(-9.75)	(-3.63)	(-3.62)
AGE		-0.00291*	-0.00287*	-0.00436**	-0.00414**	-0.00271	-0.00267	-0.00233	-0.00296	-0.00297
		(-2.11)	(-2.08)	(-3.13)	(-2.97)	(-1.74)	(-1.72)	(-1.48)	(-1.88)	(-1.89)
MIG			0.374***	0.373***	0.361***	0.346***	0.338***	0.345***	0.332***	0.327***
			(6.67)	(6.60)	(6.38)	(6.06)	(5.78)	(6.00)	(5.77)	(5.68)
FSIZE(Reference: GE 2000)										
[1] LT 20				0.719***	0.711***	0.519***	0.520***	0.511***	0.456***	0.460***
				(15.07)	(14.67)	(10.38)	(10.37)	(10.19)	(9.01)	(9.07)
[2] GE 20 LT 200				0.385***	0.378***	0.196***	0.199***	0.187***	0.150**	0.149**
				(7.82)	(7.62)	(3.84)	(3.88)	(3.65)	(2.90)	(2.88)
[3] GE 200 LT 2000				0.149**	0.162**	0.0410	0.0418	0.0352	0.0210	0.0190
				(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						-0.278***	-0.278***	-0.318***	-0.345***	-0.339***
						(-4.67)	(-4.66)	(-5.29)	(-5.73)	(-5.61)
Job With New Employer						0.129*	0.128*	0.149**	0.136*	0.135*
						(2.39)	(2.37)	(2.73)	(2.49)	(2.48)
Company Taken Over						-1.527***	-1.526***	-1.529***	-1.550***	-1.555***
						(-10.17)	(-10.16)	(-10.17)	(-10.30)	(-10.33)
Changed Job, Same Firm						-1.671***	-1.672***	-1.680***	-1.669***	-1.670***
						(-15.09)	(-15.09)	(-15.15)	(-15.04)	(-15.05)
STATE							v			
Survey year t								v	v	v
ISEI									-0.0102***	-0.0103***
									(-7.94)	(-7.99)
FEMALE										-0.0745*
										(-2.10)
Constant	0.580***	0.671***	0.556***	-0.0279	-0.215*	-0.0652	-0.0266	-0.0602	0.0143	0.0662
	(7.64)	(7.69)	(6.26)	(-0.28)	(-1.97)	(-0.58)	(-0.18)	(-0.50)	(0.12)	(0.53)
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

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

Table 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. 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[4]. 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 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 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^i_N), employed and well matched in the original occupation (m^i_N) or mismatched and employed in another occupation (x^i_N). The same holds for foreign type i individuals with corresponding notation u^i_F, m^i_F and x^i_F, so we get:

$$u^i_N + m^i_N + x^i_N = 1 - h \qquad u^i_F + m^i_F + x^i_F = h$$

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

$$e^i_N = m^i_N + x^i_N \qquad e^i_F = m^i_F + x^i_F$$

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

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

the empirical evidence that social networks generate occupational mismatch more frequently than formal search. We verify this assumption for Germany in section 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.

3.1 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 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:

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

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

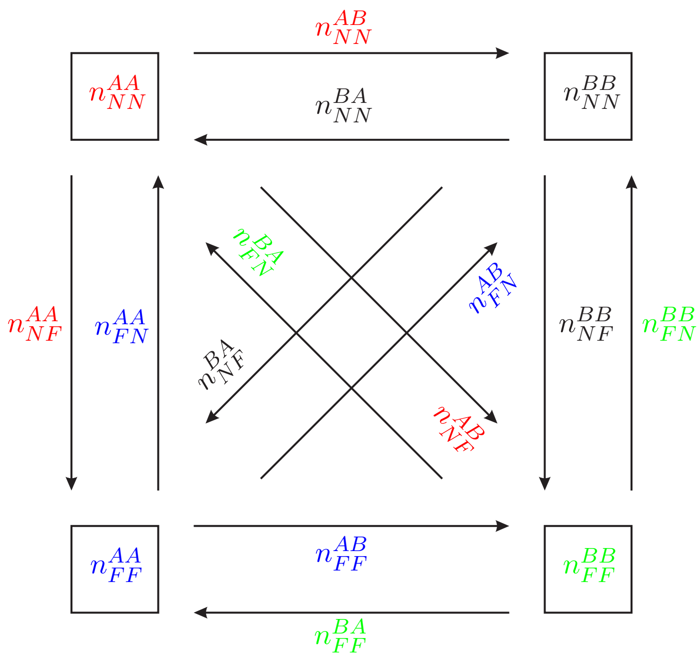

Figure 1: Composition of social networks

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:

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

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:

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

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 $(1-h)n^{AA}_{NF}$ should be the same as the total number of links from foreign individuals of type A to natives of type A given by hn^{AA}_{FN}. Moreover, the total number of links from native individuals of type B to foreign individuals of type A, that is $(1-h)n^{BA}_{NF}$, should be the same as the number of links from foreign individuals of type A to native individuals of type B given by hn^{AB}_{FN}. This means:

$$(1-h)n^{AA}_{NF} = hn^{AA}_{FN} \qquad (1-h)n^{BA}_{NF} = hn^{AB}_{FN}$$

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

$$n_{NF}^{AA} = \frac{h(1-\tau)\gamma n}{1-h} \qquad n_{NN}^{AA} = \frac{(1-2h+h\tau)\gamma n}{1-h}$$
$$n_{NF}^{BA} = \frac{h(1-\tau)(1-\gamma)n}{1-h} \qquad n_{NN}^{BA} = \frac{(1-2h+h\tau)(1-\gamma)n}{1-h}$$

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:

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

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 $\tau = 0.6$.

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.

3.2 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 2. The network job-finding rates are illustrated by the dashed arrows.

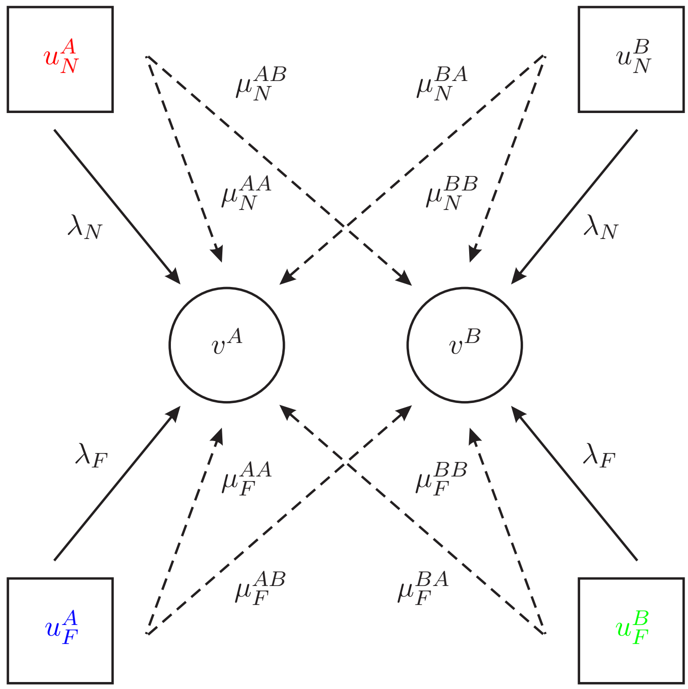

Figure 2: Structure of the labour market

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 $\mu^{AA} \equiv \mu_N^{AA}{=}\mu_F^{AA}$ 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 $\mu^{AB} \equiv \mu^{AB}_N = \mu^{AB}_F$ and:*

$$\mu^{AB} = \frac{sv_B}{u^A_N + u^A_F}\left[\frac{e^B_N}{1-h}\right]^{(1-h)\gamma n}\left[\frac{e^B_F}{h}\right]^{h\gamma n}\left(1 - \left[\frac{e^A_N}{1-h}\right]^{(1-h)(1-\gamma)n}\left[\frac{e^A_F}{h}\right]^{h(1-\gamma)n}\right)$$

The same is true in occupation B, that is $\mu^{BA}_N = \mu^{BA}_F$.

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 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^A_N and matched employment m^A_N for native type A workers can be written as:

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

Here $\delta_N(1-h-u^A_N)$ 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 $u^A_N(\lambda_N + \mu^{AA}_N + \mu^{AB}_N)$ 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 $\mu^{AA}_N + \mu^{AB}_N$. In the second equation the term $(\lambda_N + \mu^{AA}_N)u^A_N$ corresponds to native type A workers finding jobs in their primary occupation, while $\delta_N m^A_N$ is the outflow of workers from this group due to job losses.

We have two similar equations for foreign workers:

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

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 $\dot{u}^A_N = 0$, $\dot{m}^A_N = 0$, $\dot{u}^A_F = 0$ and $\dot{m}^A_F = 0$. So the steady-state distributions of workers across the three states are given by:

$$\begin{aligned}
u^A_F &= \frac{\delta_F h}{\delta_F + \lambda_F + \mu^{AA}_F + \mu^{AB}_F} \qquad m^A_F = \frac{(\lambda_F + \mu^{AA}_F)h}{\delta_F + \lambda_F + \mu^{AA}_F + \mu^{AB}_F} \qquad x^A_F = h - u^A_F - m^A_F \\
u^A_N &= \frac{\delta_N(1-h)}{\delta_N + \lambda_N + \mu^{AA}_N + \mu^{AB}_N} \qquad m^A_N = \frac{(\lambda_N + \mu^{AA}_N)(1-h)}{\delta_N + \lambda_N + \mu^{AA}_N + \mu^{AB}_N} \qquad x^A_N = 1 - h - u^A_N - m^A_N
\end{aligned} \tag{1}$$

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 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 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.

3.4 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 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 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 (1) for the equilibrium unemployment rates we find values of parameters $\{\lambda_N, \lambda_F, sv, \tau\}$ summarized in table 5.

Table 5: Exogenous parameters and target variables

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

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 6.

Table 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

Table 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 ($\tau > h$). 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.

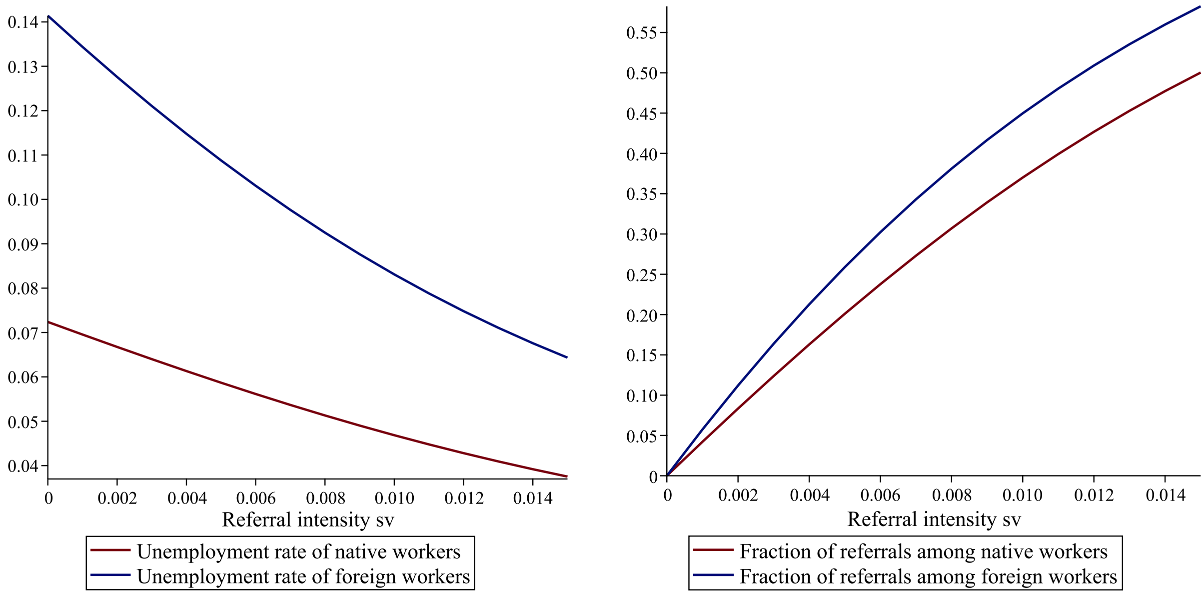

Figure 3: Left panel: Unemployment rates of native and foreign workers ($u_N/(1-h)$ and u_F/h) in the benchmark setting. Right panel: Fractions of network hires for native and foreign workers (R_N and R_F) in the benchmark setting

Our results are presented on figure 3. 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.

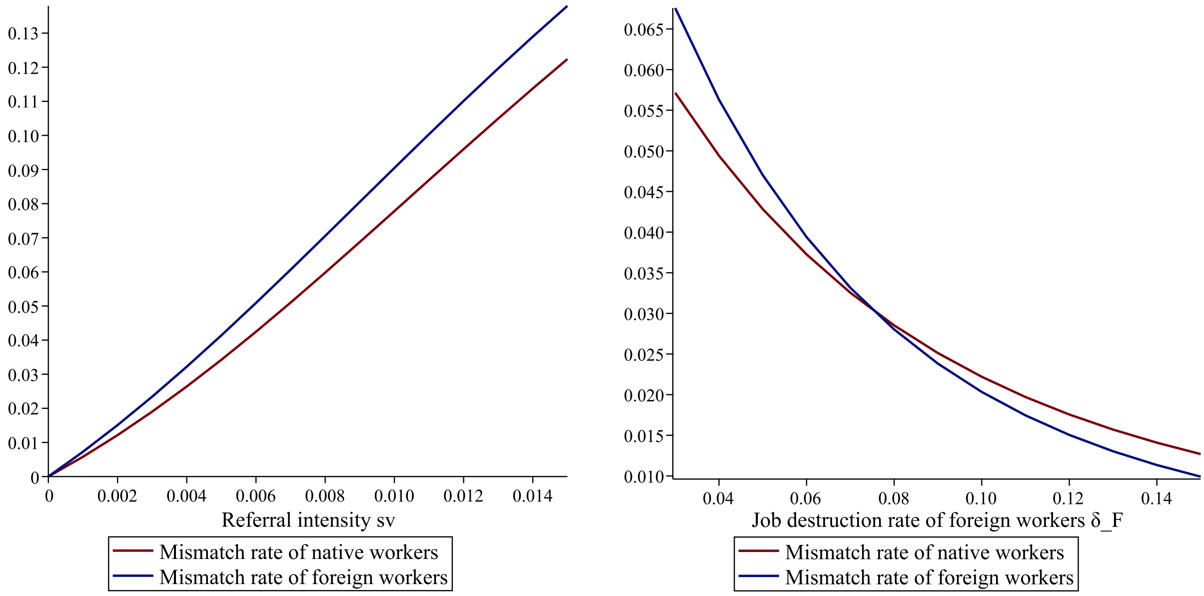

Figure 4: Left panel: Mismatch rates of native and foreign workers ($x_N^A/(1-h)$ and x_F^A/h), benchmark. Right panel: Mismatch rates of native and foreign workers ($x_N^A/(1-h)$ and x_F^A/h) for different values of δ_F

The left panel of figure 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 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 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 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.

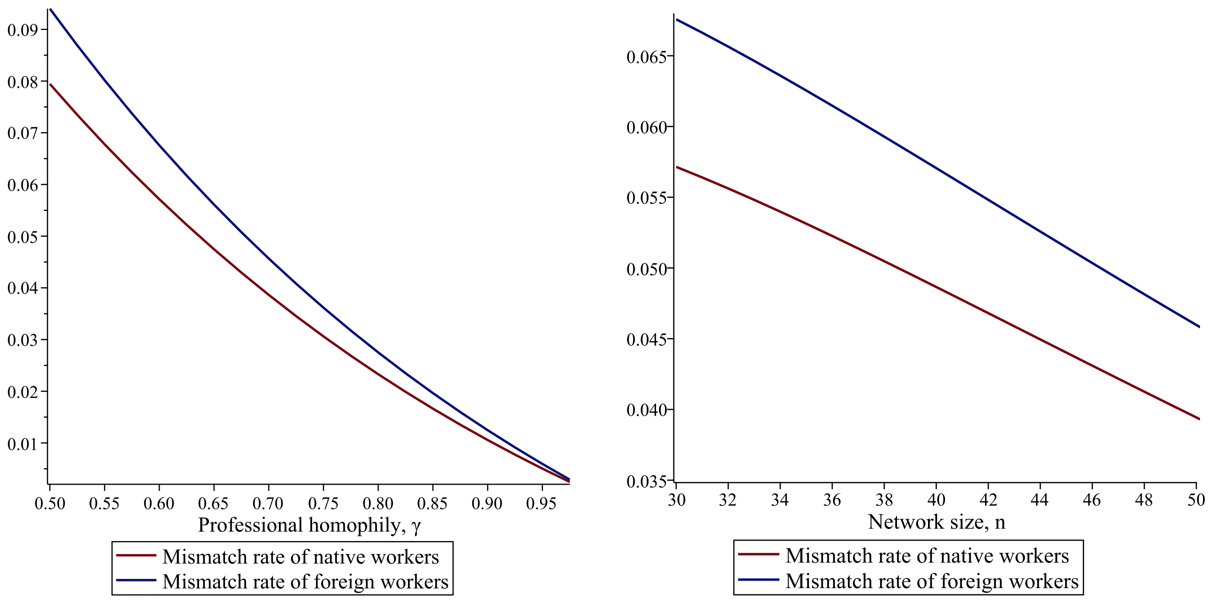

Figure 5: Left panel: Mismatch rates of native and foreign workers for different values of γ. Right panel: Mismatch rates of native and foreign workers for different values of n

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).

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?

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 7 below is slightly different from the descriptive statistics presented in Table 3.

Descriptive statistics presented in Table 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 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 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 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 15 in Appendix III. Besides statistics about the overall sample, Table 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 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 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
MATCH									
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 9[5]. 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

[5]Table 16 presented in Appendix IV includes all the coefficients of control variables.

after a long break are often associated with mismatch.

Table 9: Estimation results of occupational mismatch.

Variables	Dependent variable: MATCH								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EDU	0.230***	0.229***	0.217***	0.207***	0.210***	0.209***	0.209***	0.110***	0.110***
	(33.32)	(33.05)	(29.97)	(28.08)	(28.23)	(27.87)	(27.72)	(12.48)	(12.48)
AGE		-0.0251***	-0.0258***	-0.0233***	-0.0236***	-0.0234***	-0.0240***	-0.0240***	-0.0240***
		(-15.85)	(-16.04)	(-13.91)	(-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				-0.432***	-0.440***	-0.442***	-0.433***	-0.370***	-0.370***
				(-4.60)	(-4.68)	(-4.69)	(-4.57)	(-3.83)	(-3.82)
Job With New Employer				-0.179*	-0.180*	-0.194*	-0.204*	-0.171	-0.171
				(-1.97)	(-1.98)	(-2.13)	(-2.22)	(-1.83)	(-1.83)
Company Taken Over				0.761***	0.773***	0.772***	0.768***	0.822***	0.821***
				(5.40)	(5.48)	(5.46)	(5.42)	(5.72)	(5.71)
Changed Job, Same Firm				0.163	0.200	0.193	0.197	0.150	0.149
				(1.54)	(1.87)	(1.80)	(1.83)	(1.37)	(1.36)
FSIZE(Reference: GE 2000)									
[1] LT 20					0.155**	0.173**	0.184***	0.359***	0.360***
					(2.92)	(3.24)	(3.42)	(6.52)	(6.52)
[2] GE 20 LT 200					0.0266	0.0462	0.0569	0.166**	0.166**
					(0.50)	(0.86)	(1.06)	(3.04)	(3.04)
[3] GE 200 LT 2000					0.00808	0.0214	0.0250	0.0676	0.0676
					(0.14)	(0.37)	(0.43)	(1.14)	(1.13)
STATE						v	v	v	v
Survey year t							v	v	v
ISEI								0.0299***	0.0299***
								(21.02)	(21.02)
FEMALE									-0.00608
									(-0.16)
Constant	-2.649***	-1.723***	-1.430***	-1.210***	-1.306***	-1.228***	-1.326***	-1.583***	-1.579***
	(-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

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

In the next step, $MIG_{i,t}$ is added to the regression equation. The coefficients are presented in Table 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. 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.

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

Variables	Dependent variable: MATCH					
	(1)	(2)	(3)	(4)	(5)	(6)
MIG		-0.400***	-0.375***	-0.372***	-0.365***	0.0144
		[-0.099]	[-0.093]	[-0.092]	[-0.090]	[0.003]
REF			-0.422***	-0.421***		
			[-0.103]	[-0.103]		
MIG × REF				-0.00707		
				[-0.002]		
CHAN (Reference: Internet)						
Public emp. agency					-0.325***	-0.316***
					[-0.078]	[-0.076]
Private emp. agency					-0.258	-0.252
					[-0.062]	[-0.060]

Continued on next page

Variables	Dependent variable: MATCH					
	(1)	(2)	(3)	(4)	(5)	(6)
Newspaper					-0.235**	-0.233**
					[-0.056]	[-0.056]
Referrals					-0.530***	-0.513***
					[-0.129]	[-0.125]
Other					0.017	0.0428
					[0.004]	[0.010]
MIG × CHAN(Reference: MIG × Internet)						
MIG × Public empl. agency						-0.296
						[-0.072]
MIG × Private empl. agency						-0.267
						[-0.065]
MIG × Newspaper						-0.196
						[-0.048]
MIG × Referrals						-0.397
						[-0.097]
MIG × Other						-0.551
						[-0.134]
Control variables	v	v	v	v	v	v
Time FE	v	v	v	v	v	v
Constant	-1.583***	-1.480***	-1.285***	-1.285***	-1.123***	-1.136***
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

Marginal effects are in squared brackets. 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 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.[6]

In Table 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

[6]Table 17 presented in Appendix V includes all the coefficients of control variables.

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 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		-0.351***	-0.324***	-0.321***	-0.314***
		[-0.087]	[-0.080]	[-0.079]	[-0.077]
Indirect migrant		-0.120	-0.110	-0.122	-0.106
		[-0.029]	[-0.027]	[-0.030]	[-0.026]
Referrals			-0.420***	-0.421***	
			[-0.103]	[-0.103]	
MIGBACK × REF (Reference: German national × Formal channels)					
Direct migrant × referrals				-0.00752	
				[-0.002]	
Indirect migrant × referrals				0.0370	
				[0.009]	
Chan (Reference: Internet)					
Public emp. agency					-0.329***
					[-0.079]
Private emp. agency					-0.255
					[-0.061]
Newspaper					-0.243**
					[-0.058]
Referrals					-0.534***
					[-0.130]
Other					0.00896
					[0.002]
Control variables	v	v	v	v	v
Time FE	v	v	v	v	v
Constant	-1.583***	-1.485***	-1.293***	-1.293***	-1.125***
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

Marginal effects are 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 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.[7]

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

[7]Table 18 presented in Appendix VI includes all the coefficients of control variables.

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 $MATCH_{i,t}$[8]. This decomposition is based on the following equation:

$$\hat{Y}_N - \hat{Y}_F = \underbrace{(\bar{X}_N - \bar{X}_F)'\hat{B}_N}_{\text{Endowment effect (explained)}} + \underbrace{\bar{X}_F'(\hat{B}_N - \hat{B}_F)}_{\text{Coefficient effect (unexplained)}} \quad (2)$$

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 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 12.

Table 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 7. The decomposition shows that the endowment effect is equal to

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

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 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).

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.

5 Conclusion

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.

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).

References

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.

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.

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

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.

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.

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.

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.

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

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.

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.

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.

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

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.

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.

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

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

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.

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.

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.

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:

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

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

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

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

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

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

Appendix I. Estimation results: employment rates.

Table 13: Estimation results of employment rates, full table.

Variables	Dependent variable: EMP							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EDU	0.315***	0.314***	0.315***	0.324***	0.371***	0.371***	0.368***	0.344***
	(76.14)	(76.09)	(76.17)	(76.99)	(80.91)	(80.81)	(80.55)	(75.31)
AGE		0.00225**	0.00225**	-0.0110***	-0.00943***	-0.00971***	-0.0171***	-0.0186***
		(3.24)	(3.25)	(-13.46)	(-11.28)	(-11.56)	(-18.52)	(-20.06)
FEMALE			-0.138***	-0.107***	-0.113***	-0.115***	-0.122***	-0.134***
			(-8.71)	(-6.59)	(-6.86)	(-6.98)	(-7.37)	(-8.09)
MARST(Reference: Married)								

Continued on next page

Variables	Dependent variable: EMP							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
[2] Single				-0.867***	-0.769***	-0.774***	-0.957***	-1.003***
				(-39.95)	(-34.43)	(-34.54)	(-39.67)	(-41.38)
[3] Widowed				-0.323***	-0.300***	-0.298***	-0.346***	-0.370***
				(-5.60)	(-5.15)	(-5.12)	(-5.93)	(-6.33)
[4] Divorced				-0.770***	-0.743***	-0.747***	-0.807***	-0.839***
				(-32.41)	(-30.85)	(-30.94)	(-33.04)	(-34.18)
[5] Separated				-0.752***	-0.748***	-0.748***	-0.801***	-0.815***
				(-16.98)	(-16.62)	(-16.60)	(-17.70)	(-17.94)
Number of							-0.175***	-0.167***
Children in HH							(-20.25)	(-19.22)
Foreign								-0.652***
citizen								(-24.68)
STATE (Reference: Schleswig-Holstein)								
[2] Hamburg					0.0557	0.0557	0.0475	0.129
					(0.64)	(0.64)	(0.54)	(1.46)
[3] Lower Saxony					0.202***	0.201***	0.208***	0.218***
					(3.73)	(3.71)	(3.82)	(4.00)
[4] Bremen					-0.437***	-0.437***	-0.454***	-0.434***
					(-4.65)	(-4.65)	(-4.83)	(-4.60)
[5] North-Rhine					0.0425	0.0429	0.0376	0.0980*
-Westfalia					(0.86)	(0.87)	(0.76)	(1.98)
[6] Hessen					0.267***	0.268***	0.258***	0.337***
					(4.63)	(4.65)	(4.47)	(5.81)
[7] Rheinland-Pfalz					0.328***	0.329***	0.334***	0.359***
					(5.35)	(5.37)	(5.43)	(5.83)
[8] Baden					0.586***	0.588***	0.591***	0.719***
-Wuerttemberg					(10.83)	(10.85)	(10.91)	(13.14)
[9] Bavaria					0.618***	0.618***	0.609***	0.648***
					(11.77)	(11.76)	(11.58)	(12.30)
[10] Saarland					0.259**	0.259**	0.228*	0.267**
					(2.82)	(2.82)	(2.48)	(2.89)
[11] Berlin					-0.599***	-0.598***	-0.603***	-0.555***
					(-10.26)	(-10.24)	(-10.28)	(-9.45)
[12] Brandenburg					-0.903***	-0.903***	-0.916***	-0.933***
					(-16.34)	(-16.34)	(-16.53)	(-16.84)
[13] Mecklenburg					-0.854***	-0.854***	-0.870***	-0.894***
-Vorpommern					(-13.97)	(-13.96)	(-14.17)	(-14.57)
[14] Saxony					-0.678***	-0.679***	-0.700***	-0.722***
					(-12.95)	(-12.95)	(-13.34)	(-13.74)
[15] Saxony-Anhalt					-0.939***	-0.941***	-0.963***	-0.985***
					(-17.05)	(-17.07)	(-17.44)	(-17.84)
[16] Thuringia					-0.717***	-0.717***	-0.741***	-0.766***
					(-12.87)	(-12.87)	(-13.27)	(-13.70)
Survey year t (Reference: 2000)								
2001						-0.00268	0.000226	-0.000699
						(-0.06)	(0.01)	(-0.02)
2002						-0.0289	-0.0264	-0.0326
						(-0.65)	(-0.59)	(-0.73)
2003						-0.149***	-0.145**	-0.153***
						(-3.39)	(-3.28)	(-3.45)
2004						-0.194***	-0.190***	-0.198***
						(-4.40)	(-4.32)	(-4.49)
2005						-0.174***	-0.172***	-0.179***
						(-3.87)	(-3.83)	(-3.96)
2006						-0.187***	-0.185***	-0.198***
						(-4.23)	(-4.19)	(-4.46)
2007						0.00179	0.00295	-0.00986
						(0.04)	(0.06)	(-0.21)
2008						0.129**	0.132**	0.118*

Continued on next page

Variables	Dependent variable: EMP							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
						(2.65)	(2.69)	(2.42)
2009						0.0803	0.0861	0.0713
						(1.69)	(1.81)	(1.49)
2010						-0.146***	-0.0740	-0.0926*
						(-3.44)	(-1.74)	(-2.17)
2011						0.0131	0.0887*	0.0681
						(0.30)	(2.06)	(1.57)
2012						0.0223	0.0986*	0.0796
						(0.52)	(2.28)	(1.84)
2013						-0.0788	-0.00783	0.0347
						(-1.94)	(-0.19)	(0.85)
2014						-0.0157	0.0551	0.0868*
						(-0.37)	(1.28)	(2.02)
Constant	-1.311***	-1.406***	-1.349***	-0.542***	-1.115***	-1.050***	-0.520***	-0.125
	(-28.11)	(-25.53)	(-24.25)	(-9.04)	(-14.31)	(-12.72)	(-6.01)	(-1.42)
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

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

Appendix II. Estimation results: referral hiring.

Table 14: Estimation results of referral hiring, full table.

Variables	Dependent variable: REF									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
EDU	-0.106***	-0.106***	-0.0988***	-0.0798***	-0.0707***	-0.0641***	-0.0643***	-0.0649***	-0.0291***	-0.0290***
	(-17.48)	(-17.33)	(-16.05)	(-12.68)	(-10.80)	(-9.67)	(-9.62)	(-9.75)	(-3.63)	(-3.62)
AGE		-0.00291*	-0.00287*	-0.00436**	-0.00414**	-0.00271	-0.00267	-0.00233	-0.00296	-0.00297
		(-2.11)	(-2.08)	(-3.13)	(-2.97)	(-1.74)	(-1.72)	(-1.48)	(-1.88)	(-1.89)
Foreign citizen			0.374***	0.373***	0.361***	0.346***	0.338***	0.345***	0.332***	0.327***
			(6.67)	(6.60)	(6.38)	(6.06)	(5.78)	(6.00)	(5.77)	(5.68)
FSIZE(Reference: GE 2000)										
[1] LT 20				0.719***	0.711***	0.519***	0.520***	0.511***	0.456***	0.460***
				(15.07)	(14.67)	(10.38)	(10.37)	(10.19)	(9.01)	(9.07)
[2] GE 20 LT 200				0.385***	0.378***	0.196***	0.199***	0.187***	0.150**	0.149**
				(7.82)	(7.62)	(3.84)	(3.88)	(3.65)	(2.90)	(2.88)
[3] GE 200 LT 2000				0.149**	0.162**	0.0410	0.0418	0.0352	0.0210	0.0190
				(2.67)	(2.89)	(0.71)	(0.73)	(0.61)	(0.36)	(0.33)
IND(Reference: Services)										
[1] Agriculture					-0.0130	-0.0125	0.00331	0.00233	-0.102	-0.128
					(-0.10)	(-0.10)	(0.03)	(0.02)	(-0.80)	(-1.01)
[2] Energy					-0.357	-0.300	-0.299	-0.273	-0.254	-0.274
					(-1.78)	(-1.46)	(-1.46)	(-1.33)	(-1.24)	(-1.33)
[3] Mining					0.420	0.552	0.538	0.515	0.490	0.451
					(1.03)	(1.31)	(1.27)	(1.21)	(1.15)	(1.06)
[4] Manufacturing					0.152**	0.120*	0.119*	0.111*	0.0997*	0.0751
					(3.12)	(2.43)	(2.40)	(2.23)	(2.01)	(1.47)
[5] Construction					0.0173	-0.0130	-0.00487	-0.0157	-0.0414	-0.0818
					(0.33)	(-0.24)	(-0.09)	(-0.30)	(-0.77)	(-1.44)
[6] Trade					0.226***	0.184***	0.186***	0.181***	0.191***	0.187***
					(5.18)	(4.18)	(4.21)	(4.10)	(4.30)	(4.22)
[7] Transport					0.168*	0.138	0.142*	0.128	0.111	0.0816

Continued on next page

Variables	Dependent variable: REF									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					(2.39)	(1.93)	(1.98)	(1.78)	(1.54)	(1.11)
[8] Bank,Insurance					-0.205*	-0.0971	-0.0978	-0.111	-0.0447	-0.0530
					(-2.03)	(-0.94)	(-0.94)	(-1.07)	(-0.43)	(-0.51)
TOJCH(Reference: First job)										
Job After						-0.278***	-0.278***	-0.318***	-0.345***	-0.339***
Break						(-4.67)	(-4.66)	(-5.29)	(-5.73)	(-5.61)
Job With New						0.129*	0.128*	0.149**	0.136*	0.135*
Employer						(2.39)	(2.37)	(2.73)	(2.49)	(2.48)
Company						-1.527***	-1.526***	-1.529***	-1.550***	-1.555***
Taken Over						(-10.17)	(-10.16)	(-10.17)	(-10.30)	(-10.33)
Changed Job,						-1.671***	-1.672***	-1.680***	-1.669***	-1.670***
Same Firm						(-15.09)	(-15.09)	(-15.15)	(-15.04)	(-15.05)
STATE(Reference: Schleswig-Holstein)										
[2] Hamburg							-0.00578			
							(-0.04)			
[3] Lower							-0.0863			
Saxony							(-0.80)			
[4] Bremen							-0.0680			
							(-0.33)			
[5] North-Rhine							-0.000841			
-Westfalia							(-0.01)			
[6] Hessen							-0.0279			
							(-0.25)			
[7] Rheinland-							-0.0504			
Pfalz							(-0.42)			
[8] Baden-							-0.0547			
Wuerttemberg							(-0.52)			
[9] Bavaria							-0.0266			
							(-0.26)			
[10] Saarland							0.0321			
							(0.18)			
[11] Berlin							-0.00456			
							(-0.04)			
[12]Brandenburg							-0.0533			
							(-0.44)			
[13]Mecklenburg							-0.318*			
-Vorpommern							(-2.24)			
[14] Saxony							-0.110			
							(-0.98)			
[15] Saxony-							-0.0240			
Anhalt							(-0.20)			
[16] Thuringia							0.0714			
							(0.59)			
Survey year t (Reference: 2000)										
2001								-0.0160	-0.0135	-0.0131
								(-0.22)	(-0.18)	(-0.18)
2002								-0.0761	-0.0720	-0.0700
								(-0.98)	(-0.92)	(-0.90)
2003								0.0891	0.0956	0.0961
								(1.11)	(1.19)	(1.20)
2004								-0.0186	-0.0178	-0.0200
								(-0.23)	(-0.22)	(-0.25)
2005								0.0636	0.0653	0.0628
								(0.78)	(0.80)	(0.77)
2006								0.00971	0.00428	0.00188
								(0.12)	(0.05)	(0.02)
2007								0.210**	0.205**	0.205**
								(2.71)	(2.65)	(2.65)

Continued on next page

Variables	Dependent variable: REF									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
2008								0.150	0.145	0.145
								(1.92)	(1.85)	(1.85)
2009								0.160*	0.163*	0.161*
								(2.06)	(2.09)	(2.07)
2010								-0.218**	-0.224**	-0.222**
								(-2.58)	(-2.65)	(-2.62)
2011								-0.150	-0.154	-0.151
								(-1.73)	(-1.76)	(-1.73)
2012								-0.119	-0.132	-0.130
								(-1.33)	(-1.47)	(-1.44)
2013								-0.163	-0.159	-0.156
								(-1.81)	(-1.76)	(-1.73)
2014								-0.0529	-0.0680	-0.0595
								(-0.69)	(-0.89)	(-0.77)
ISEI									-0.0102***	-0.0103***
									(-7.94)	(-7.99)
FEMALE										-0.0745*
										(-2.10)
Constant	0.580***	0.671***	0.556***	-0.0279	-0.215*	-0.0652	-0.0266	-0.0602	0.0143	0.0662
	(7.64)	(7.69)	(6.26)	(-0.28)	(-1.97)	(-0.58)	(-0.18)	(-0.50)	(0.12)	(0.53)
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

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

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

Table 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%
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						

Continued on next page

	German citizens	Foreign citizens	German nationals	Direct migrants	Indirect migrants	Overall
[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 16: Estimation results of occupational mismatch.

Variables	Dependent variable: MATCH								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EDU	0.230***	0.229***	0.217***	0.207***	0.210***	0.209***	0.209***	0.110***	0.110***
	(33.32)	(33.05)	(29.97)	(28.08)	(28.23)	(27.87)	(27.72)	(12.48)	(12.48)
AGE		-0.0251***	-0.0258***	-0.0233***	-0.0236***	-0.0234***	-0.0240***	-0.0240***	-0.0240***
		(-15.85)	(-16.04)	(-13.91)	(-14.07)	(-13.90)	(-14.15)	(-13.95)	(-13.95)
IND(Reference: Services)									
Agriculture			-0.223	-0.179	-0.207	-0.155	-0.152	0.136	0.133
			(-1.65)	(-1.31)	(-1.51)	(-1.12)	(-1.10)	(0.96)	(0.94)
Energy			-0.450**	-0.512**	-0.475**	-0.469**	-0.481**	-0.571**	-0.572**
			(-2.59)	(-2.92)	(-2.70)	(-2.66)	(-2.73)	(-3.21)	(-3.21)
Mining			-0.932	-0.993*	-0.960*	-0.878	-0.874	-0.857	-0.860

Continued on next page

Variables	Dependent variable: MATCH								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
			(-1.96)	(-2.05)	(-1.99)	(-1.82)	(-1.81)	(-1.76)	(-1.77)
Manufacturing			-0.208***	-0.220***	-0.201***	-0.213***	-0.202***	-0.177***	-0.179**
			(-4.03)	(-4.23)	(-3.83)	(-4.04)	(-3.82)	(-3.30)	(-3.24)
Construction			0.337***	0.343***	0.341***	0.360***	0.369***	0.443***	0.439***
			(5.94)	(6.00)	(5.98)	(6.25)	(6.40)	(7.53)	(7.03)
Trade			-0.418***	-0.398***	-0.409***	-0.415***	-0.409***	-0.442***	-0.442***
			(-8.67)	(-8.18)	(-8.37)	(-8.49)	(-8.35)	(-8.93)	(-8.93)
Transport			-0.862***	-0.888***	-0.876***	-0.878***	-0.870***	-0.819***	-0.821***
			(-11.41)	(-11.64)	(-11.44)	(-11.44)	(-11.33)	(-10.58)	(-10.44)
Bank,Insurance			0.234*	0.168	0.195	0.172	0.183	-0.0103	-0.0108
			(2.30)	(1.63)	(1.88)	(1.65)	(1.76)	(-0.10)	(-0.10)
TOJCH(Reference: First job)									
Job After Break				-0.432***	-0.440***	-0.442***	-0.433***	-0.370***	-0.370***
				(-4.60)	(-4.68)	(-4.69)	(-4.57)	(-3.83)	(-3.82)
Job With New				-0.179*	-0.180*	-0.194*	-0.204*	-0.171	-0.171
Employer				(-1.97)	(-1.98)	(-2.13)	(-2.22)	(-1.83)	(-1.83)
Company Taken				0.761***	0.773***	0.772***	0.768***	0.822***	0.821***
Over				(5.40)	(5.48)	(5.46)	(5.42)	(5.72)	(5.71)
Changed Job,				0.163	0.200	0.193	0.197	0.150	0.149
Same Firm				(1.54)	(1.87)	(1.80)	(1.83)	(1.37)	(1.36)
FSIZE(Reference: GE 2000)									
[1] LT 20					0.155**	0.173**	0.184***	0.359***	0.360***
					(2.92)	(3.24)	(3.42)	(6.52)	(6.52)
[2] GE 20 LT 200					0.0266	0.0462	0.0569	0.166**	0.166**
					(0.50)	(0.86)	(1.06)	(3.04)	(3.04)
[3] GE 200					0.00808	0.0214	0.0250	0.0676	0.0676
LT 2000					(0.14)	(0.37)	(0.43)	(1.14)	(1.13)
STATE(Reference: Bavaria)									
[1] Schleswig-						-0.114	-0.117	-0.0471	-0.0473
Holstein						(-1.03)	(-1.05)	(-0.42)	(-0.42)
[2] Hamburg						0.0222	0.0158	-0.00693	-0.00733
						(0.15)	(0.11)	(-0.05)	(-0.05)
[3] Lower Saxony						-0.0886	-0.0881	-0.0378	-0.0379
						(-1.24)	(-1.23)	(-0.52)	(-0.52)
[4] Bremen						-0.125	-0.118	-0.135	-0.136
						(-0.64)	(-0.61)	(-0.69)	(-0.69)
[5] North-Rhine-						-0.0487	-0.0410	0.00297	0.00275
Westfalia						(-0.81)	(-0.68)	(0.05)	(0.05)
[6] Hessen						0.0820	0.0809	0.0849	0.0846
						(1.05)	(1.03)	(1.07)	(1.07)
[7] Rheinland-						-0.136	-0.145	-0.0889	-0.0888
Pfalz						(-1.47)	(-1.55)	(-0.94)	(-0.94)
[8] Baden-						-0.0221	-0.0270	-0.0151	-0.0152
Wuerttemberg						(-0.33)	(-0.40)	(-0.22)	(-0.22)
[10] Saarland						-0.327	-0.321	-0.313	-0.314
						(-1.84)	(-1.81)	(-1.74)	(-1.74)
[11] Berlin						-0.155	-0.164	-0.119	-0.119
						(-1.56)	(-1.65)	(-1.18)	(-1.18)
[12] Brandenburg						-0.374***	-0.365***	-0.265**	-0.265**
						(-4.07)	(-3.97)	(-2.84)	(-2.84)
[13]Mecklenburg-						-0.317**	-0.311**	-0.243*	-0.244*
Vorpommern						(-2.80)	(-2.74)	(-2.11)	(-2.12)
[14] Saxony						-0.0300	-0.0331	0.0512	0.0508
						(-0.39)	(-0.43)	(0.65)	(0.65)
[15] Saxony-						-0.202*	-0.194*	-0.0988	-0.0994
Anhalt						(-2.21)	(-2.13)	(-1.06)	(-1.07)
[16] Thuringia						-0.266**	-0.264**	-0.140	-0.140
						(-2.88)	(-2.85)	(-1.48)	(-1.48)
Survey year t (Reference: 2000)									

Continued on next page

Variables	Dependent variable: MATCH								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
2001							0.0702	0.0617	0.0616
							(0.88)	(0.76)	(0.76)
2002							0.157	0.151	0.151
							(1.86)	(1.76)	(1.76)
2003							0.0845	0.0647	0.0648
							(0.96)	(0.72)	(0.73)
2004							0.0848	0.0933	0.0931
							(0.97)	(1.06)	(1.05)
2005							0.145	0.130	0.129
							(1.62)	(1.43)	(1.43)
2006							0.0291	0.0287	0.0285
							(0.34)	(0.33)	(0.33)
2007							0.0721	0.0899	0.0900
							(0.86)	(1.05)	(1.06)
2008							0.118	0.133	0.133
							(1.38)	(1.54)	(1.54)
2009							0.238**	0.232**	0.232**
							(2.79)	(2.69)	(2.69)
2010							0.0437	0.0676	0.0679
							(0.49)	(0.74)	(0.74)
2011							0.132	0.143	0.144
							(1.43)	(1.53)	(1.54)
2012							0.239*	0.283**	0.284**
							(2.49)	(2.91)	(2.91)
2013							0.249**	0.233*	0.234*
							(2.60)	(2.39)	(2.40)
2014							0.268**	0.317***	0.318***
							(3.21)	(3.74)	(3.75)
ISEI								0.0299***	0.0299***
								(21.02)	(21.02)
FEMALE									-0.00608
									(-0.16)
Constant	-2.649***	-1.723***	-1.430***	-1.210***	-1.306***	-1.228***	-1.326***	-1.583***	-1.579***
	(-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

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

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

Table 17: Estimation results of occupational mismatch using citizenship and search channels.

Variables	Dependent variable: MATCH					
	(1)	(2)	(3)	(4)	(5)	(6)
EDU	0.110***	0.108***	0.106***	0.106***	0.103***	0.103***
	(0.027)	(0.026)	(0.026)	(0.026)	(0.025)	(0.025)
AGE	-0.0240***	-0.0243***	-0.0246***	-0.0246***	-0.0240***	-0.0240***
	(-0.006)	(-0.006)	(-0.006)	(-0.006)	(-0.006)	(-0.006)
IND(Reference: Services)						
[1] Agriculture	0.136	0.145	0.141	0.141	0.133	0.138
	(0.032)	(0.034)	(0.033)	(0.033)	(0.031)	(0.032)

Continued on next page

Variables	Dependent variable: MATCH					
	(1)	(2)	(3)	(4)	(5)	(6)
[2] Energy	-0.571**	-0.580**	-0.602***	-0.601***	-0.602***	-0.599***
	(-0.141)	(-0.143)	(-0.149)	(-0.149)	(-0.149)	(-0.148)
[3] Mining	-0.857	-0.842	-0.821	-0.821	-0.852	-0.850
	(-0.211)	(-0.207)	(-0.202)	(-0.202)	(-0.210)	(-0.209)
[4] Manufacturing	-0.177***	-0.172**	-0.166**	-0.166**	-0.165**	-0.163**
	(-0.043)	(-0.042)	(-0.040)	(-0.040)	(-0.040)	(-0.039)
[5] Construction	0.443***	0.448***	0.445***	0.445***	0.440***	0.439***
	(0.099)	(0.101)	(0.100)	(0.100)	(0.099)	(0.099)
[6] Trade	-0.442***	-0.439***	-0.424***	-0.424***	-0.432***	-0.431***
	(-0.109)	(-0.108)	(-0.104)	(-0.104)	(-0.106)	(-0.106)
[7] Transport	-0.819***	-0.807***	-0.804***	-0.804***	-0.812***	-0.812***
	(-0.202)	(-0.199)	(-0.198)	(-0.198)	(-0.200)	(-0.200)
[8] Bank,Insurance	-0.0103	-0.0140	-0.0253	-0.0253	-0.0260	-0.0258
	(-0.002)	(-0.003)	(-0.006)	(-0.006)	(-0.006)	(-0.006)
TOJCH(Reference: First job)						
Job After Break	-0.370***	-0.377***	-0.418***	-0.418***	-0.437***	-0.436***
	(-0.090)	(-0.092)	(-0.102)	(-0.102)	(-0.106)	(-0.106)
Job With New Employer	-0.171	-0.180	-0.171	-0.171	-0.170	-0.170
	(-0.041)	(-0.043)	(-0.041)	(-0.041)	(-0.040)	(-0.040)
Company Taken Over	0.822***	0.815***	0.710***	0.710***	0.644***	0.646***
	(0.168)	(0.167)	(0.147)	(0.147)	(0.134)	(0.135)
Changed Job, Same Firm	0.150	0.133	0.0345	0.0346	-0.0659	-0.0687
	(0.035)	(0.031)	(0.008)	(0.008)	(-0.015)	(-0.016)
FSIZE(Reference: GE 2000)						
[1] LT 20	0.359***	0.361***	0.406***	0.406***	0.423***	0.425***
	(0.087)	(0.088)	(0.098)	(0.098)	(0.102)	(0.103)
[2] GE 20 LT 200	0.166**	0.166**	0.182***	0.182***	0.201***	0.203***
	(0.041)	(0.041)	(0.045)	(0.045)	(0.050)	(0.050)
[3] GE 200 LT 2000	0.0676	0.0682	0.0676	0.0676	0.0808	0.0821
	(0.017)	(0.017)	(0.017)	(0.017)	(0.020)	(0.020)
STATE(Reference: Bavaria)						
[1] Schleswig-Holstein	-0.0471	-0.0741	-0.0647	-0.0647	-0.0658	-0.0669
	(-0.011)	(-0.018)	(-0.016)	(-0.016)	(-0.016)	(-0.016)
[2] Hamburg	-0.00693	-0.0251	-0.0281	-0.0281	-0.0290	-0.0279
	(-0.002)	(-0.006)	(-0.007)	(-0.007)	(-0.007)	(-0.007)
[3] Lower Saxony	-0.0378	-0.0550	-0.0577	-0.0577	-0.0623	-0.0619
	(-0.009)	(-0.013)	(-0.014)	(-0.014)	(-0.015)	(-0.015)
[4] Bremen	-0.135	-0.148	-0.151	-0.151	-0.141	-0.145
	(-0.033)	(-0.036)	(-0.037)	(-0.037)	(-0.034)	(-0.035)
[5] North-Rhine-Westfalia	0.00297	0.000882	0.00462	0.00462	0.000943	0.000663
	(0.001)	(0.000)	(0.001)	(0.001)	(0.002)	(0.000)
[6] Hessen	0.0849	0.0845	0.0848	0.0849	0.0739	0.0737
	(0.020)	(0.020)	(0.020)	(0.020)	(0.018)	(0.018)
[7] Rheinland-Pfalz	-0.0889	-0.0996	-0.0971	-0.0971	-0.102	-0.103
	(-0.021)	(-0.024)	(-0.024)	(-0.024)	(-0.025)	(-0.025)
[8] Baden-Wuerttemberg	-0.0151	0.00362	-0.00112	-0.00113	-0.00236	-0.00215
	(-0.004)	(0.001)	(-0.000)	(-0.000)	(-0.001)	(-0.000)
[10] Saarland	-0.313	-0.304	-0.297	-0.297	-0.305	-0.298
	(-0.077)	(-0.075)	(-0.073)	(-0.073)	(-0.075)	(-0.073)
[11] Berlin	-0.119	-0.120	-0.117	-0.117	-0.119	-0.119
	(-0.029)	(-0.029)	(-0.028)	(-0.028)	(-0.029)	(-0.029)
[12] Brandenburg	-0.265**	-0.299**	-0.301**	-0.301**	-0.297**	-0.298**
	(-0.065)	(-0.074)	(-0.074)	(-0.074)	(-0.073)	(-0.073)
[13] Mecklenburg-Vorpommern	-0.243*	-0.280*	-0.292*	-0.292*	-0.283*	-0.284*
	(-0.060)	(-0.069)	(-0.072)	(-0.072)	(-0.069)	(-0.070)
[14] Saxony	0.0512	0.0165	0.0111	0.0111	0.0119	0.0106
	(0.012)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)
[15] Saxony-Anhalt	-0.0988	-0.136	-0.140	-0.140	-0.128	-0.129
	(-0.024)	(-0.033)	(-0.034)	(-0.034)	(-0.031)	(-0.031)

Continued on next page

Variables	Dependent variable: MATCH					
	(1)	(2)	(3)	(4)	(5)	(6)
[16] Thuringia	-0.140	-0.177	-0.163	-0.163	-0.158	-0.159
	(-0.034)	(-0.043)	(-0.040)	(-0.040)	(-0.039)	(-0.039)
Survey year t (Reference: 2000)						
2001	0.0617	0.0576	0.0519	0.0519	0.0537	0.0511
	(0.015)	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)
2002	0.151	0.147	0.138	0.138	0.127	0.125
	(0.037)	(0.036)	(0.034)	(0.034)	(0.031)	(0.030)
2003	0.0647	0.0564	0.0571	0.0571	0.0533	0.0520
	(0.016)	(0.014)	(0.014)	(0.014)	(0.013)	(0.013)
2004	0.0933	0.0822	0.0844	0.0844	0.0704	0.0695
	(0.023)	(0.020)	(0.020)	(0.020)	(0.017)	(0.017)
2005	0.130	0.117	0.119	0.119	0.101	0.100
	(0.032)	(0.029)	(0.029)	(0.029)	(0.025)	(0.025)
2006	0.0287	0.0177	0.0137	0.0137	-0.00251	-0.00315
	(0.007)	(0.004)	(0.003)	(0.003)	(-0.001)	(-0.001)
2007	0.0899	0.0795	0.0947	0.0946	0.0746	0.0727
	(0.022)	(0.019)	(0.023)	(0.023)	(0.018)	(0.018)
2008	0.133	0.121	0.130	0.130	0.109	0.107
	(0.033)	(0.030)	(0.032)	(0.032)	(0.027)	(0.026)
2009	0.232**	0.226**	0.236**	0.236**	0.215*	0.213*
	(0.057)	(0.055)	(0.057)	(0.057)	(0.052)	(0.052)
2010	0.0676	0.0579	0.0296	0.0296	0.00378	0.00391
	(0.017)	(0.014)	(0.007)	(0.007)	(0.001)	(0.001)
2011	0.143	0.127	0.103	0.103	0.0666	0.0633
	(0.035)	(0.031)	(0.025)	(0.025)	(0.016)	(0.0015)
2012	0.283**	0.267**	0.249*	0.249*	0.212*	0.214*
	(0.069)	(0.065)	(0.060)	(0.060)	(0.051)	(0.052)
2013	0.233*	0.218*	0.201*	0.201*	0.165	0.165
	(0.057)	(0.053)	(0.049)	(0.049)	(0.040)	(0.040)
2014	0.317***	0.318***	0.310***	0.310***	0.269**	0.265**
	(0.077)	(0.077)	(0.075)	(0.075)	(0.065)	(0.064)
ISEI	0.0299***	0.0296***	0.0289***	0.0289***	0.0284***	0.0284***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Foreign		-0.400***	-0.375***	-0.372***	-0.365***	0.0144
citizen		(-0.099)	(-0.093)	(-0.092)	(-0.090)	(0.003)
Referrals			-0.422***	-0.421***		
			(-0.103)	(-0.103)		
Foreign citizen ×				-0.00707		
Referrals				(-0.002)		
CHAN (Reference: Internet)						
Public emp. agency					-0.325***	-0.316***
					(-0.078)	(-0.076)
Private emp. agency					-0.258	-0.252
					(-0.062)	(-0.060)
Newspaper					-0.235**	-0.233**
					(-0.056)	(-0.056)
Referrals					-0.530***	-0.513***
					(-0.129)	(-0.125)
Other					0.017	0.0428
					(0.004)	(0.010)
MIG × CHAN(Reference: Foreign citizen × Internet)						
Foreign citizen × Public						-0.296
emp. agency						(-0.072)
Foreign citizen × Private						-0.267
emp. agency						(-0.065)
Foreign citizen ×						-0.196
Newspaper						(-0.048)
Foreign citizen × Referrals						-0.397
						(-0.097)

Continued on next page

Variables	Dependent variable: MATCH					
	(1)	(2)	(3)	(4)	(5)	(6)
Foreign citizen × Other						-0.551
						(-0.134)
Constant	-1.583***	-1.480***	-1.285***	-1.285***	-1.123***	-1.136***
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

Marginal effects are 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 18: Estimation results of occupational mismatch using migration background and search channels.

Variables	Dependent variable: MATCH				
	(1)	(2)	(3)	(4)	(5)
EDU	0.110***	0.109***	0.107***	0.107***	0.104***
	(0.027)	(0.026)	(0.026)	(0.026)	(0.025)
AGE	-0.0240***	-0.0240***	-0.0243***	-0.0243***	-0.0238***
	(-0.006)	(-0.006)	(-0.006)	(-0.006)	(-0.006)
IND(Reference: Services)					
[1] Agriculture	0.136	0.138	0.135	0.135	0.127
	(0.032)	(0.032)	(0.032)	(0.032)	(0.030)
[2] Energy	-0.571**	-0.575**	-0.596***	-0.596***	-0.597***
	(-0.141)	(-0.142)	(-0.147)	(-0.147)	(-0.148)
[3] Mining	-0.857	-0.857	-0.836	-0.836	-0.866
	(-0.211)	(-0.211)	(-0.206)	(-0.206)	(-0.213)
[4] Manufacturing	-0.177***	-0.168**	-0.162**	-0.162**	-0.161**
	(-0.043)	(-0.041)	(-0.039)	(-0.039)	(-0.039)
[5] Construction	0.443***	0.449***	0.446***	0.446***	0.441***
	(0.099)	(0.101)	(0.100)	(0.100)	(0.099)
[6] Trade	-0.442***	-0.438***	-0.423***	-0.423***	-0.431***
	(-0.109)	(-0.108)	(-0.104)	(-0.104)	(-0.106)
[7] Transport	-0.819***	-0.806***	-0.803***	-0.803***	-0.811***
	(-0.202)	(-0.199)	(-0.198)	(-0.198)	(-0.200)
[8] Bank,Insurance	-0.0103	-0.0215	-0.0323	-0.0322	-0.0328
	(-0.002)	(-0.005)	(-0.008)	(-0.008)	(-0.008)
TOJCH(Reference: First job)					
Job After Break	-0.370***	-0.379***	-0.419***	-0.419***	-0.437***
	(-0.090)	(-0.092)	(-0.102)	(-0.102)	(-0.106)
Job With New Employer	-0.171	-0.184	-0.174	-0.174	-0.173
	(-0.041)	(-0.044)	(-0.042)	(-0.042)	(-0.041)
Company Taken Over	0.822***	0.811***	0.707***	0.707***	0.644***
	(0.168)	(0.166)	(0.147)	(0.147)	(0.134)
Changed Job, Same Firm	0.150	0.129	0.0319	0.0319	-0.0657
	(0.035)	(0.030)	(0.007)	(0.007)	(-0.015)
FSIZE(Reference: GE 2000)					
[1] LT 20	0.359***	0.360***	0.405***	0.405***	0.421***
	(0.087)	(0.087)	(0.098)	(0.098)	(0.102)
[2] GE 20 LT 200	0.166**	0.169**	0.185***	0.185***	0.204***
	(0.041)	(0.042)	(0.046)	(0.046)	(0.050)
[3] GE 200 LT 2000	0.0676	0.0687	0.0681	0.0683	0.0812
	(0.017)	(0.017)	(0.017)	(0.017)	(0.020)

Continued on next page

Variables	Dependent variable: MATCH				
	(1)	(2)	(3)	(4)	(5)
STATE(Reference: Bavaria)					
[1] Schleswig-Holstein	-0.0471	-0.0673	-0.0578	-0.0577	-0.0593
	(-0.011)	(-0.016)	(-0.014)	(-0.014)	(-0.014)
[2] Hamburg	-0.00693	-0.0173	-0.0207	-0.0198	-0.0220
	(-0.002)	(-0.004)	(-0.005)	(-0.005)	(-0.005)
[3] Lower Saxony	-0.0378	-0.0427	-0.0461	-0.0461	-0.0509
	(-0.009)	(-0.010)	(-0.011)	(-0.011)	(-0.012)
[4] Bremen	-0.135	-0.113	-0.119	-0.120	-0.110
	(-0.033)	(-0.027)	(-0.029)	(-0.029)	(-0.027)
[5] North-Rhine-Westfalia	0.00297	0.0110	0.0140	0.0141	0.0100
	(0.001)	(0.003)	(0.003)	(0.003)	(0.002)
[6] Hessen	0.0849	0.0939	0.0932	0.0932	0.0819
	(0.020)	(0.022)	(0.022)	(0.022)	(0.020)
[7] Rheinland-Pfalz	-0.0889	-0.0824	-0.0813	-0.0813	-0.0866
	(-0.022)	(-0.020)	(-0.020)	(-0.020)	(-0.021)
[8] Baden-Wuerttemberg	-0.0151	0.00864	0.00349	0.00373	0.00191
	(-0.004)	(0.002)	(0.001)	(0.001)	(0.000)
[10] Saarland	-0.313	-0.280	-0.274	-0.273	-0.282
	(-0.077)	(-0.069)	(-0.067)	(-0.067)	(-0.069)
[11] Berlin	-0.119	-0.116	-0.114	-0.114	-0.116
	(-0.029)	(-0.028)	(-0.028)	(-0.028)	(-0.028)
[12] Brandenburg	-0.265**	-0.305**	-0.306**	-0.306**	-0.302**
	(-0.065)	(-0.075)	(-0.075)	(-0.075)	(-0.074)
[13] Mecklenburg-Vorpommern	-0.243*	-0.289*	-0.300**	-0.300**	-0.290*
	(-0.060)	(-0.071)	(-0.074)	(-0.074)	(-0.071)
[14] Saxony	0.0512	0.00759	0.00316	0.00324	0.00440
	(0.012)	(0.002)	(0.001)	(0.001)	(0.001)
[15] Saxony-Anhalt	-0.0988	-0.143	-0.146	-0.146	-0.134
	(-0.024)	(-0.035)	(-0.038)	(-0.036)	(-0.033)
[16] Thuringia	-0.140	-0.190*	-0.174	-0.174	-0.169
	(-0.034)	(-0.047)	(-0.043)	(-0.043)	(-0.041)
Survey year t (Reference: 2000)					
2001	0.0617	0.0611	0.0549	0.0548	0.0565
	(0.015)	(0.015)	(0.013)	(0.013)	(0.014)
2002	0.151	0.146	0.138	0.138	0.126
	(0.037)	(0.036)	(0.034)	(0.034)	(0.031)
2003	0.0647	0.0591	0.0594	0.0592	0.0555
	(0.016)	(0.015)	(0.015)	(0.015)	(0.014)
2004	0.0933	0.0922	0.0932	0.0933	0.0788
	(0.023)	(0.023)	(0.023)	(0.023)	(0.019)
2005	0.130	0.124	0.125	0.124	0.106
	(0.032)	(0.030)	(0.031)	(0.030)	(0.026)
2006	0.0287	0.0259	0.0209	0.0207	0.00420
	(0.007)	(0.006)	(0.005)	(0.005)	(0.001)
2007	0.0899	0.0894	0.104	0.103	0.0831
	(0.022)	(0.022)	(0.025)	(0.025)	(0.020)
2008	0.133	0.133	0.141	0.141	0.119
	(0.033)	(0.033)	(0.035)	(0.034)	(0.029)
2009	0.232**	0.238**	0.246**	0.246**	0.224*
	(0.057)	(0.058)	(0.060)	(0.060)	(0.054)
2010	0.0676	0.0682	0.0393	0.0393	0.0127
	(0.017)	(0.017)	(0.010)	(0.010)	(0.003)
2011	0.143	0.143	0.118	0.118	0.0813
	(0.035)	(0.035)	(0.029)	(0.029)	(0.020)
2012	0.283**	0.278**	0.258**	0.258**	0.221*
	(0.069)	(0.067)	(0.063)	(0.063)	(0.054)
2013	0.233*	0.229*	0.212*	0.212*	0.175
	(0.057)	(0.056)	(0.052)	(0.052)	(0.043)
2014	0.317***	0.340***	0.330***	0.330***	0.288***

Continued on next page

Variables	Dependent variable: MATCH				
	(1)	(2)	(3)	(4)	(5)
	(0.077)	(0.082)	(0.079)	(0.079)	(0.069)
ISEI	0.0299***	0.0293***	0.0286***	0.0286***	0.0281***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
MIGBACK (Reference: German national)					
Direct migrant		-0.351***	-0.324***	-0.321***	-0.314***
		(-0.087)	(-0.080)	(-0.079)	(-0.077)
Indirect migrant		-0.120	-0.110	-0.122	-0.106
		(-0.029)	(-0.027)	(-0.030)	(-0.026)
Referrals			-0.420***	-0.421***	
			(-0.103)	(-0.103)	
MIGBACK × REF (Reference: German national × Formal channels)					
Direct migrant × referrals				-0.00752	
				(-0.002)	
Indirect migrant × referrals				0.0370	
				(0.009)	
Chan (Reference: Internet)					
Public emp. agency					-0.329***
					(-0.079)
Private emp. agency					-0.255
					(-0.061)
Newspaper					-0.243**
					(-0.058)
Referrals					-0.534***
					(-0.130)
Other					0.00896
					(0.002)
Constant	-1.583***	-1.485***	-1.293***	-1.293***	-1.125***
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

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