

On the location choice of newly arrived immigrants in Germany*

Carsten Pohl†

September 18, 2007

Abstract

The aging and the shrinkage of the population in Germany will come along with (regional) shortages in the labor force. Since immigration might mitigate the decline of the labor supply I study the initial location choice of newly arrived foreign-born using micro data from the German Socio-Economic Panel (1991-2004) as well as macro data from official statistics. As a main result, I find that migration networks in Western Germany have a significant impact on the probability of incoming foreign-born originating in the same source country to move to these ethnic clusters. In contrast, in Eastern Germany only concrete social ties between former and newly arrived foreign-born positively influence the probability of incoming immigrants to locate in the same region.

Keywords: Immigrants, Location Choice, Migration Networks, Discrete Choice Models

JEL classification: C25, F22, J61

*I thank Chad Carroll, Hien Truong and seminar participants at the Dresden University of Technology for helpful comments and discussion.

†ifo Institute for Economic Research, Branch Dresden, Einsteinstrasse 3, 01069 Dresden, Germany, Email: pohl@ifo.de, phone: +49 351 2647624

1 Introduction

In most industrialized economies, immigrants tend to be clustered in specific regions within a country. For instance, in the United States almost two thirds of all legal immigrants are concentrated in five states [Rytina (2006)]. In Germany, there are significant differences in the regional distribution of the foreign-born as well. Over 60 % of all foreigners live in three West German states but only 6 % reside in East Germany (without Berlin) [Federal Statistical Office of Germany (2007)]. Whereas empirical studies for other industrialized countries show that migration networks seem to play an important role in the location choice of new immigrants there is only little evidence for Germany.

However, the analysis of networks effects on the location choice of incoming immigrants is important since Germany is exposed to demographic changes. According to population projections, the aging and the decline of the labor force will proceed much quicker in Germany than in other industrialized countries. In addition, there are significant differences on a regional level indicating that East Germany will be more affected by demographic changes than West Germany. Furthermore, rural areas will face a stronger decline in the working population than urban areas. Overall, the aging and the shrinkage of the labor force in Germany will come along with shortages especially for high-skilled workers [Grundig, Pohl, and Thum (2007)]. Consequently, demographic development will entail negative effects on (regional) economic growth if vacant positions cannot be filled. In this context, the immigration of qualified foreign-born workers might mitigate regional shortages in the German labor market.

Against these stylized facts, I analyze the location choice of newly arrived immigrants in Germany. For this purpose, I use macro data on the regional level as well as micro data on individuals for the period 1991-2004. Since the installed basis of foreign-born in a region as well as social ties between former and new immigrants are supposed to influence the initial location choice of incoming immigrants to a large extent I investigate whether migration networks are important in the location decision. As a main result, I find that ethnic clusters in Western Germany have a significant impact on the probability of incoming foreign-born originating in the same source country to move to these migration networks. In contrast, in Eastern Germany only concrete social ties between former and newly arrived foreign-born

positively influence the probability of incoming immigrants to locate in the same region.

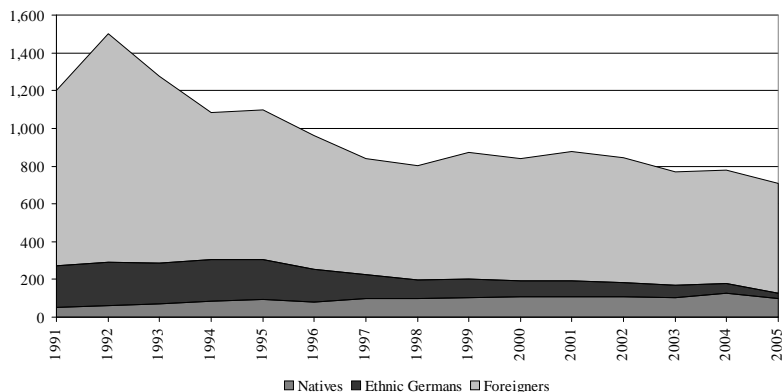
The remainder of this paper is structured as follows. In section 2, I provide some additional background information on the composition of the migration inflow to Germany as well as the regional distribution of immigrants in Germany. Section 3 reviews the related literature on the location choice of incoming immigrants. Section 4 presents the used micro and macro data sets. Section 5 outlines the econometric specifications that capture individual-specific characteristics of the immigrants as well as alternative-specific data on the destination regions. In section 6 empirical results are presented and discussed. Section 7 summarizes the findings and concludes with regard to policy implications.

2 Migration flows and the regional distribution of immigrants in Germany

With the fall of the iron curtain at the beginning of the 1990s a new wave of immigrants moved to Germany. This inflow consisted of immigrants born in Eastern European Countries or the former Soviet Union. Immigrants from these regions have been ethnic Germans ("Aussiedler") who are legally considered as Germans since they have German ancestors and "usual" foreign-born. In addition to this inflow, family-related migration from former guest-worker countries continued to represent the most important share of the immigrant inflow. As Figure 1 reveals, immigration to Germany has consisted mainly of foreign-born individuals between 1991 and 2005 (on average 77 %). During this period ethnic and native Germans have only represented 13 % and 10 % of the gross immigrant inflow, respectively.

With regard to the share of immigrants on the total population there are large differences between federal states. The share of foreigners ranges between 1.9 to 2.8 % in East Germany (without Berlin). In West Germany (without the city states Bremen and Hamburg), the share of immigrants is significantly higher amounting up to 11.9 % in Baden-Wuerttemberg. The regional distribution of migrants also shows that only around 6 % of all foreigners live in Eastern Germany (without Berlin) although 18 % of the native population are residents in this part of Germany. Table 1 provides evidence that there are differences in the regional

Figure 1: Composition of gross immigrant inflow in thousand persons (1991-2005)



distribution of foreigners when controlling for the country of origin. For instance, immigrants from Russia are less concentrated compared to immigrants from guest-workers countries such as Turkey. Almost 75 % of the Turkish foreigners in Germany are living in Baden-Wurttemberg, Bavaria, Hesse and Northrhine-Westphalia. In contrast, only 56 % of the Russian immigrants as well as of the native population have settled in these four states.

The composition of the immigration inflow as well as the regional distribution of foreigners in Germany represent the starting point for my empirical investigation. I will analyze the impact of migration networks on the location choice of incoming immigrants in Germany. This is particularly important since the total labor force in Germany will significantly decrease within the near future resulting in (regional) bottlenecks for qualified workers. Since macro data on regions as well as micro data on individuals are available I will show to what extent location-specific attributes as well as socio-economic characteristics of the immigrants are associated with initial location choice decision.

3 Related literature

There is an extensive literature on the cross-border migration decision of individuals [e.g. Borjas (1987)]. In general, it is assumed that an individual pursues the maximization of his life time utility. Thereby, migration will take place because the expected and discounted life time utility in the host country subtracted by migration costs exceeds the net present value

Table 1: Regional distribution of natives and migrants in Germany (2005)

Federal State	Share of foreigners	Regional distribution of...			
		Natives	Foreigners	Turks	Russians
Baden-Wuerttemberg (W)	11.9	12.5	17.6	17.0	11.5
Bavaria (W)	9.5	15.0	16.1	13.2	14.4
Berlin (E)	13.7	3.9	6.2	6.9	7.3
Brandenburg (E)	2.6	3.3	0.9	0.1	2.4
Bremen (W)	12.7	0.8	1.2	1.7	1.5
Hamburg (W)	14.2	2.0	3.4	3.1	4.1
Hesse (W)	11.4	7.2	9.5	10.7	7.4
Lower Saxony (W)	6.7	9.9	7.4	6.3	9.8
Mecklenb. West. Pomerania (E)	2.3	2.2	0.5	0.1	1.9
Northrhine-Westphalia (W)	10.7	21.4	26.7	33.5	22.3
Rhineland Palatinate (W)	7.7	5.0	4.3	4.0	4.7
Saarland (W)	8.3	1.3	1.2	0.8	1.4
Saxony (E)	2.8	5.6	2.1	0.2	4.0
Saxony-Anhalt (E)	1.9	3.6	1.6	0.1	2.4
Schleswig-Holstein (W)	5.4	3.3	0.6	2.0	3.2
Thuringia (E)	2.0	3.1	0.7	0.1	1.8
Total	8.9	100	100	100	100

E = East Germany, W = West Germany

Source: Federal Statistical Office of Germany (2006).

of life time utility in the home country. In this context, earnings from labor usually serve as an approximation for utility so that the size of the earnings differential between source and destination country is supposed to be decisive for the migration decision [Sjaastad (1962)].

Following this rationale, an immigrant will choose the region in the destination country which yields to the highest utility level. In this context, migration networks are supposed to provide a value added to new immigrants eventually determining their initial location choice to a large extent. Migration networks are usually described as ethnic clusters in the host country. Due to personal contacts between former immigrants and individuals originating

in the same country new immigrants are attracted to these clusters. Living in an ethnic neighborhood in the destination country is associated with lower entry barriers to the labour market and facilitated access to housing [Borjas (1995)].

In a broader definition, migration networks are also considered as ethnic clusters in the destination country. However, in contrast to the previous description there do not exist any social ties between former and new migrants prior to the location choice. But, new immigrants are supposed to move to ethnic clusters because immigrants assume to overcome teething troubles in the host country more easily compared to the situation when they would live on their own. The availability of ethnic goods serves as an additional explanation why newly arrived immigrants settle in these clusters without any social ties prior to their immigration.¹

The location choice of immigrants has been studied extensively for the United States while there is only little empirical evidence for European countries. Bartel (1989) analyzes the post-1964 location choices of immigrants in the U.S. showing that immigrants are more geographically clustered than natives. In addition, a location pattern subject to the educational attainment of the foreign-born is discovered: the higher the educational achievement the less the propensity for ethnic concentration. Zavodny (1999) confirms a positive relationship between the location choice of incoming immigrants and the installed basis of foreign-born. Controlling for the wage level as well as for the labor market situation in the destination region there only seems to exist a small positive impact on the initial location choice of immigrants. Scott, Coomes, and Izyumov (2005) explicitly focus on employment-based immigration and show that these type of foreign-born are moving to areas with a higher wage level.² Kaushal (2005) studies the interdependence between welfare benefits and the location choice of newly arrived immigrants in the United States. Overall, he only finds weak evidence that immigrants are positively affected by access to federal benefits in

¹Migration networks may not only exist in the destination but also in the sending country. In this context, Winters, de Janvry, and Sadoulet (2001) find that community and family networks in the country of origin are substitutes in the migration process. If migration is well established in a community, family networks, i.e. family-related contacts to the host country, become less important.

²Munshi (2003) investigates the network effect for Mexican immigrants in the U.S. labor market. He finds that the existence and the size of an ethnic cluster significantly improves the labour market outcome of the foreign-born.

their location choice.

Bauer, Epstein, and Gang (2002) analyze the relationship between the stock of immigrants in a particular region in the United States and the arrival of new immigrants. They do not find a linear but a hump-shaped relationship. Immigrants seem to move to a particular region as long as the positive effects associated with the migration network exceed the negative effects resulting from increased competition on the regional labor market among the immigrant population.

In contrast to the extensive literature for the United States there are only few studies for European countries. Edin and Aslund (2003) and Aslund (2005) show that immigrants in Sweden are living in ethnic enclaves since it significantly improves their labor market outcomes. Damm (2007) studies the determinants of the immigrants' location choices in Denmark showing that newly arrived foreign-born settle in locations where previous immigrants live.

With regard to the location choice of immigrants in Germany only the study by Bauer and Zimmermann (1997) provides some insights. Focusing on the role of migration networks within the group of ethnic Germans ("Aussiedler") and individuals from the German Democratic Republic ("Übersiedler") they show that these two groups of immigrants settle close to relatives and friends at the beginning of their stay in West Germany. However, these social ties seem to become less important with increasing migration duration.³

The study by Bauer and Zimmermann (1997) is limited in several aspects in order to draw general inferences on the location choice of incoming migrants. First, only ethnic Germans as well as citizens from the GDR were captured in this analysis. However, as Figure 1 shows native as well as ethnic Germans only represent a small fraction of the overall immigration inflow in the 1990s. Second, the focus in their study is not on the initial location choice but on the existence of social ties between immigrants. They consider individuals who have been living in Germany up to ten years showing differences in social contacts between short- and long-term residents. Third, the investigation of these social contacts are limited to the three best friends of an immigrant. In particular, it is analyzed whether among these friends

³Dietz (1999) as well as Haug and Sauer (2006) present descriptive statistics that ethnic Germans are clustered in specific regions in Germany.

one person is originating from the same country. However, it is not clear whether this best friend lives in the neighborhood, in another city in Germany or even abroad. Hence, the empirical findings on the importance of migration networks for foreign-born in Germany may be subject to a significant bias.

Due to these shortcomings I extend the analysis considering the location choice of newly arrived foreign-born in Germany during the period 1991-2004. With regard to the importance of migration networks the micro data that I use provides information whether new immigrants had social ties to Germany prior to their immigration. In addition, I do not only take socio-economic characteristics of the immigrants into account but also attributes of the destination regions using a macro data set. Hence, combining micro and macro data I offer a deeper understanding to what extent location factors and immigrants characteristics are associated with the location decision.

4 Data

For the empirical investigation, I use the German Socio-Economic Panel (GSOEP), data from the Federal Statistical Office as well data from the Federal Criminal Police Office of Germany. The GSOEP is a nationally representative panel data set which was established in 1984.⁴ It is conducted on an annual basis providing information on various socio-economic characteristics of the interviewed individuals. Since I am particularly interested in the location choice of immigrants after the fall of the iron curtain when unified Germany experienced a large inflow of foreign-born I confine my analysis using data for the period 1991-2004. I do not only analyze the location choice of ethnic Germans but that of all foreign-born except asylum seekers.⁵ Only foreign-born individuals who have moved to Germany for the first time are considered.

Overall, 1 ,031 individuals aged 18 and above have immigrated to Germany between 1991-2004. Due to the relatively small number of observations on a regional level I aggregate

⁴For detailed information on data access and availability of the German Socio-Economic Panel (GSOEP) see Haisken-DeNew and Frick (2005).

⁵In Germany, the regional distribution of asylum seekers is subject to agreements between the federal states. Hence, they cannot be assumed to freely decide on their initial location choice.

Table 2: Aggregation of German regions

Region	Federal States
East Germany	Berlin, Brandenburg, Mecklenburg Western Pomerania, Saxony, Saxony-Anhalt and Thuringia
West Germany	Northrhine-Westphalia, Rhineland Palatinate and Saarland
North Germany	Bremen, Hamburg, Schleswig-Holstein and Lower Saxony
South Germany	Baden-Wuerttemberg, Bavaria and Hesse

the sixteen federal states of Germany to four regions: *East, West, North and South Germany* (see Table 2).⁶ With regard to the destination regions of the newly arrived immigrants in Germany I find that 5.2 % moved to the East, 23.7 % to the North, 34.8 % to the West and 36.3 % to the South.

Descriptive statistics on socio-demographic characteristics of the foreign-born are provided in Table 3. With respect to the educational achievements I distinguish between low- and high-skilled immigrants using the International Standard Classification of Education (ISCED).⁷ As *low-skilled* I consider individuals who belong to groups 0 to 4 of the ISCED, i.e. persons who have as a maximum a lower secondary educational degree. Immigrants who have at least an advanced secondary degree are regarded as *high-skilled* (groups 5 and 6). During 1991-2004 around 28 % of all incoming immigrants had an advanced degree. The average *age at immigration* was 33.4 years and 47% were *males*.⁸

With regard to the role of migration networks the German Socio-Economic Panel (GSOEP) provides data whether a foreign-born had contacts to individuals in Germany prior to their immigration. I find that 39 % of all newly arrived immigrants in the sample had *social*

⁶In order to test the independence of irrelevant alternatives assumption I conducted a Hausman-McFadden test [Long and Freese (2006)]. The test shows that the four aggregated regions represent a distinct set of alternatives.

⁷The ISCED comprises seven levels: pre-primary (0), primary education (1), lower secondary education (2), upper secondary education (3), post-secondary non-tertiary education (4), first stage of tertiary education (5), and second stage of tertiary education (6).

⁸I do not take the employment status of the incoming immigrants into account since no information whether incoming migrants are provided with working permits was available for the entire period.

Table 3: Descriptive statistics on newly arrived immigrants in Germany

Variable	Share
High-skilled (ISCED 5-6)	28%
Age in years (std. dev.)	33,4 (14.8)
Male	47%
Ethnic Germans	35%
Social ties prior to immigration	39%
Country/region of origin	
Turkey	13%
Eastern Europe	27%
Western Europe	8%
Russia	41%
Rest of the world	11%
Destination region	
North Germany	23.7%
West Germany	36.3%
East Germany	5.2%
South Germany	36.3%
Observations (n)	1,031

Source: GSOEP (1991-2004).

ties to inhabitants in Germany prior to immigration. Decomposing the immigrant inflow by sending region I find that 13 % of all incoming immigrants were from *Turkey* and only 8 % from *Western Europe*.⁹ The large majority originate in *Eastern Europe*¹⁰ (27 %) as well as in *Russia* (41 %). Around 35 % of incoming immigrants in the GSOEP were *ethnic Germans*. Hence, this specific group of individuals are overrepresented in the GSOEP when compared to official statistics (see 1).¹¹

⁹Due to the small number of observations on single countries in Western Europe I created the group *Western European Countries* consisting of countries that belong to the European Economic Area.

¹⁰As for Western Europe I created the group *Eastern European Countries* (Bulgaria, Estonia, Ex-Yugoslavia, Czech Republic, Hungary, Latvia, Lithuania, Poland, Romania, Slovenia and Slovakia).

¹¹Note, that a federal regulation with regard to the regional distribution of ethnic Germans was only introduced in the year 1994 [Haug and Sauer (2006)]. However, the settlement was not enforced by law and

Table 4: Descriptive statistics on macro-economic variables (1991-2004)

	Average	Std. dev.	North	South	West	East
GDP per employee (in Euro)	51,356	1,735	53,204	57,328	56,308	36,184
Unemployment rate (in %)	8,4	3,1	8.6	5.8	9.6	17.3
Crimes per 1000 inhabitants	76	17	99	65	72	103
Share foreigners (in %)	9.2	2.1	6.9	10.9	9.8	3.8
Turkey	2.7	0.7	2.1	3.3	2.8	0.8
Eastern Europe ^{a)}	4.5	1.5	3.0	6.1	4.6	1.2
Western Europe ^{b)}	0.4	0.2	0.2	0.3	0.6	0.6
Russia ^{c)}	0.1	0.1	0.1	0.1	0.2	0.3
Rest of the World ^{d)}	1.5	0.2	1.5	1.1	1,6	0.9

a) Central and Eastern European Countries. - b) Western European Countries.

c) Countries of the former Soviet Union. - d) All other countries

Source: Federal Statistical Office of Germany (2007) and Federal Criminal Police Office of Germany (2007); own calculations.

Since the micro data set does not only provide the year of immigration but also the federal states to which individuals moved into, micro and macro data from official statistics may be matched. Descriptive statistics on the four corresponding German regions are provided in Table 4.

As explanatory variables I include macro-economic variables which are supposed to have an impact on the location choice of newly arriving immigrants. First, as in previous studies I consider *GDP per employee* in the estimation [Bartel (1989)]. Thereby, South Germany represents the wealthiest region within Germany closely followed by West and North Germany. GDP per employee in East Germany only attains 70 % of the German average in the period 1991-2004.

Second, the *unemployment rate* is considered as an approximation of the regional labour market situation. Table 4 shows that the average unemployment rate was 17.3 % in East Germany during 1991-2004. Again, the South German region is an outperformer with an

ignoring the regional distribution was sanctioned by a cut in state-aid for integration. For these reasons, ethnic Germans are assumed to choose their residence area in Germany on their own.

average unemployment rate of 5.8 % only.

Third, Nechyba (1998) shows that personal security represents an important factor in the location choice as well. For this reason, I include the number of *crimes per 1,000 inhabitants* as an explanatory variable. Official statistics indicate that the average number for Germany is 76 cases.¹²

Finally, the overall share of foreigners as well as the share of specific source regions/countries of the immigrants are taken into account. The share of foreigners varies across the four aggregate areas ranging from 3.8 % in East Germany to 10.9 % in South Germany. In addition, differences in the spatial distribution of the immigrants exist with regard to their country/region of origin.

Overall, these statistics suggest that South Germany represents an attractive region with regard to economic well-being, favorable labor market situation and a low crime rate. The overall economic situation turns out to be worse in East Germany. Not only is the GDP per employee relatively low but also the unemployment rate is above the German average.

5 Econometric framework

In order to analyze the location choice of newly arrived immigrants I use the framework of a random utility model (RUM). The RUM serves as a starting point for the econometric analysis since it is possible to derive various discrete choice models as shown in McFadden (1974), McFadden (2001) and Train (2003).

First, I apply a multinomial logit model capturing case-specific variables only, i.e. characteristics of the immigrants but no data on the four regions. Second, I estimate a conditional logit model indicating to what extent the attributes of the different regions are affecting the location choice of the foreign-born. These two econometric specifications serve at the same time as robustness checks for the following model specifications. Finally, I combine

¹²Note, that no data on xenophobic crimes was available for the entire considered period (1991-2004) so that I used the overall number of crimes per region.

case- and alternative-specific characteristics arriving at mixed logit models. In these specifications, characteristics of the immigrants as well as on the German regions are taken into consideration. The simple mixed logit model assumes that there exists no correlation between the unobserved factors for the different regions. Hence, I extend the analysis estimating an alternative-specific multinomial probit model where correlations are explicitly taken into account.

A random utility model

Newly arriving immigrants in Germany are considered as utility maximizers who may choose between four different regions for their initial location. The underlying idea of random utility models is that the utility the i th individual ($i = 1, \dots, N$) derives from the j th alternative of a given choice set K ($j = 1, \dots, K$) is represented by U_{ij} . Thereby, utility depends on various determining variables represented by V . These variables can be decomposed into M factors which are specific to the individual (migrant) and into L factors which are specific to the (location) choice ($V = M + L$). Since the relationship between utility U_{ij} and the determining variables is also subject to factors not considered in V_{ij} a random disturbance term ϵ_{ij} is included. In other words, the error term captures factors that are not known to the researcher (unobservables) but that affect utility of the individual. Hence, the utility function may be written as follows:

$$U_{ij} = V_{ij} + \epsilon_{ij} \tag{1}$$

Since the choice of a specific region represents a discrete outcome non-linear or limited dependent variable models are appropriate for the investigation. These models are also labelled as discrete choice models and correspond to the theoretical framework of random utility models [Greene (2008)].

In order apply discrete choice models the decision maker (migrant) is supposed to choose among a set of alternatives (locations) that meets three criteria [Train (2003)]: (1) the number of alternatives in the set is finite; (2) the alternatives are mutually exclusive; and (3) all possible alternatives are included in the model. Since in this investigation immigrants may

choose between four different regions within Germany, the prerequisites for the application of discrete choice models are satisfied.

By assumption, an incoming migrant will choose location $j = k$ if it corresponds among all alternatives ($j = 1, \dots, K$) to the highest utility level. Hence, the probability that an individual i will choose alternative k may be written as follows if Y_i is considered as a random variable [Cameron and Trivedi (2005) and Greene (2005)]:

$$\Pr(Y_i = k) = \Pr(U_{ik} > U_{ij}) \text{ for all } j = 1, \dots, K \quad j \neq k \quad (2)$$

Since the utility U_{ij} may be decomposed into V_{ij} and ϵ_{ij} equation 2 can be rearranged arriving at following expressions:

$$\Pr(Y_i = k) = \Pr(V_{ik} + \epsilon_{ik} > V_{ij} + \epsilon_{ij}) \quad (3)$$

$$\Pr(Y_i = k) = \Pr(\epsilon_{ij} - \epsilon_{ik} < V_{ik} - V_{ij}) \quad (4)$$

In order to estimate this model an assumption about the disturbance term ϵ_{ij} has to be made. McFadden (1974) shows that if all the K disturbances are independent and identically distributed (i.i.d.) with a Gumbel distribution, then the generalized logit model is derived:

$$\Pr(Y_i = k) = \frac{\exp(V_{ik})}{\sum_{j=1}^K \exp(V_{ij})} \quad (5)$$

In equation 5 the error terms have cancelled out. This is due to the independence assumption that the unobserved portion of utility for one alternative ϵ_{ij} is not correlated with the unobserved portion for any another alternative ϵ_{ik} . Since equation 5 is relatively easy to compute with maximum likelihood technique the (generalized) logit model has become popular in the literature. Equation 5 represents the starting point for the following econometric specifications.

Multinomial and conditional logit model

In the multinomial logit model, only individual characteristics are taken into account, i.e., no data on the attributes of the location is considered. Note, that immigrant characteristics Z do only vary with the individual i but not with the region. If the value of a specific individual characteristic (Z_i) increases, the utility of individual i choosing location k will rise if $\beta_j > 0$ and fall if $\beta_j < 0$.

$$\Pr(Y_i = k) = \frac{\exp(\beta_j Z_i)}{\sum_{j=1}^K \exp(\beta_j Z_i)} \quad (6)$$

Since it is not very plausible to assume that only individual characteristics influence the immigrants' location choice I also estimate a conditional logit model. However, in this specification only state attributes W_{jl} but no characteristics of the immigrants are taken into consideration. In line with the previous model specification an increase in a particular attribute of a location (W_{jl}) will result in a higher utility for individual i if $\gamma > 0$ and in a lower utility if $\gamma < 0$.

$$\Pr(Y_i = k) = \frac{\exp(\gamma W_{jl})}{\sum_{j=1}^K (\gamma W_{jl})} \quad (7)$$

However, neither location-specific (alternative-specific) attributes nor individual (case-specific) characteristics can be assumed to determine the location choice independently.

Mixed logit models

Finally, I combine the multinomial and conditional logit model arriving at the mixed logit model.¹³ In equation 8 the location choice is supposed to be dependent on state characteristics as in the conditional logit model [see equation 7]. In addition, immigrant characteristics are incorporated in order to account for person-specific characteristics [see equation 6].

$$\Pr(Y_i = k) = \frac{\exp(\beta_j Z_i + \gamma W_{jl})}{\sum_{j=1}^K \exp(\beta_j Z_i + \gamma W_{jl})} \quad (8)$$

For estimation purposes, equation 8 has to be modified. In the mixed logit model state attributes enter as in the conditional logit model. However, since individual characteristics do not vary across the alternatives and would fall out of the probability, I incorporate interaction terms between alternatives and individual characteristics in equation 9 [Cameron and Trivedi (2005)]:

$$\Pr(Y_i = k) = \frac{\exp(\beta_j (Z_i \cdot W_{jl}) + \gamma W_{jl})}{\sum_{j=1}^K \exp(\beta_j (Z_i \cdot W_{jl}) + \gamma W_{jl})} \quad (9)$$

The mixed logit model obviously offers a more complete picture to what extent alternative- and case-specific variables are associated with the location choice of newly arrived immigrants in Germany. However, so far disturbances ϵ_{ij} are assumed to be independent and identically distributed (i.i.d.) implying that the unobserved factors are not correlated over alternatives. This assumption which is also labelled as independent from irrelevant alternatives (IIA) is restrictive since utilities for the alternatives may be correlated with each other. The so-called mixed logit model with error components which is also labelled as alternative-specific multinomial probit model (ASMNP) relaxes this restrictive structure of the covariance matrix of the error terms [Long and Freese (2006)].

¹³I follow Cameron and Trivedi (2005) who use the term mixed logit model when combining the conditional and the multinomial logit model. In addition, I also use a mixed logit model in the sense of Train (2003) where error components that create correlations among the utilities of the different alternatives are allowed.

In particular, the ASMNP assumes that errors are joint normally distributed allowing correlations of unobserved factors over alternatives. Until recently, this model used to be difficult to estimate since it is computational exhaustive to fit multidimensional normal integrals. Recent developments have improved the fitting of the model using maximum simulated likelihood technique. However, in order to identify the covariance matrix of the errors constraints have to be imposed [Train (2003)]. These constraints guarantee that neither adding a constant to the utility for each alternative (region) nor dividing each utility by a constant will affect the location choice of the immigrant.

Alternative structures for the covariance matrix of the errors are suitable. First, if the errors had a unit variance and no correlation, it would be the counterpart to the multinomial logit model with the no correlation condition. Second, it is possible to have one correlation coefficient that is common to all error pairs and allow the remaining correlations to vary over alternatives. Third, in the most general form the correlation coefficients differ for each pair of errors.¹⁴

6 Empirical results

In the following section, I will represent the empirical results of the different econometric specifications. First, evidence from the multinomial logit model indicating the relationship between socio-economic characteristics and the location choice is given. Second, results from the conditional logit model show to what extent location-specific attributes influence the propensity to locate in a specific German region. Finally, empirical evidence from the mixed logit models indicate the relationship between immigrant characteristics and regional factors on the location choice probability.

¹⁴In my empirical analysis I only obtained results for the first and second specification whereas the (simulated) maximum likelihood function did not converge for the third alternative. As Cameron and Trivedi (2005) point, out the latter specification is not easy to estimate with more than four choices and has only had little empirical success because of multidimensional integrals of the likelihood function.

Multinomial and conditional logit model

The estimated marginal effects of the multinomial logit model are provided for each destination region in Table 5.¹⁵ In Model 1 only three socio-economic characteristics, i.e. *age at immigration* (in years) and the dummy variables *male* and *high-skilled*, are included. The results indicate that females are more likely to move to East Germany whereas males seem to have a preference for the three remaining regions. With regard to age there only exists a small negative effect for the East German region indicating that older immigrants display a slightly lower probability to settle in the new federal states when entering Germany. Looking at the educational background of the immigrants I only find a negative marginal effect for the South. Lower educated immigrants prefer South Germany as initial location.

In Model 2 additional explanatory variables are considered (see Table 6). The marginal effects of this specification show again that females are more likely to move to Eastern Germany whereas males seem to prefer the Western part of Germany (North, South and West Germany). A possible explanation for the difference between East and West Germany is that male individuals are first movers and females join their husbands in East Germany. In contrast, in Western Germany predominantly male individuals settle due to the better situation on the labour market. Again, the marginal effects for the high-skilled immigrants is negative for Southern Germany implying that low-skilled immigrants are more likely to move to this area.

In order to reveal whether there are differences in the location choice by country of origin I include a set of dummy variables where *Rest of the World* serves as the reference category. I find Turkish immigrants to have a higher probability to move to South Germany whereas the marginal effect turns out to be negative for Eastern Germany. In fact, all considered immigrant groups are less likely to move to Eastern Germany. Russian and Turkish immigrants prefer South Germany as destination region since the marginal effect is positive and significant. East and West European immigrants are more likely to move to West Germany.

¹⁵Following Greene (2008) I only report standard errors for the coefficients and not for the marginal effects since inference is about the former but not the later.

Table 5: Marginal effects of the multinomial logit model (Model 1)

Model 1	East Germany	North Germany	South Germany	West Germany
	marg. eff.	marg eff.	marg eff.	marg eff.
Outcome probability	0.053	0.242	0.358	0.347
Male	-0.042**	0.030*	0.036*	0.036*
Age at immigration	-0.001**	-0.001	0.001	-0.001
High-skilled	0,008	0.006	-0.054*	0.040
Observations	1,031			
Pseudo R ²	0.1091			
Log Likelihood	-1,273			

p-value: *** = 0.01, ** = 0.05, * = 0.10

Source: own calculations based on GSOEP 1991-2004.

With regard to social ties of newly arrived immigrants I find that this explanatory variable seems to be particularly important for East Germany. Individuals who have social ties prior to immigration in East Germany display a positive marginal effect to move to this region whereas no statistically significant relationship is found for the other German areas.

Next, the results of the marginal effects of the conditional logit model are presented where only alternative-specific attributes but no socio-economic characteristics of the immigrants are taken into account. In this specification, Eastern Germany serves as the reference category since descriptive statistics as well as the results of the multinomial logit model indicate that this region seems to have a different immigration pattern.

In Model 3 only the different destination regions are included as regressors. I find that North, South and West Germany are preferred destination regions by newly arrived immigrants. In Model 4, I control for macro-economic variables, the overall share of foreigners and the share of specific source regions of immigrants. Again, the three West German regions are significantly attracting more immigrants than East Germany. The unemployment situation seems not to be a statistically significant variable. Gross domestic product per employee turns out to be positive and significant. The marginal effect of the crime rate has a negative sign and is significant. Hence, immigrants seem to prefer wealthier and safer regions when entering Germany. Interestingly, the share of foreigners does not seem to have an impact at

Table 6: Marginal effects of the multinomial logit model (Model 2)

Model 2	East Germany	North Germany	South Germany	West Germany
	marg eff.	marg eff.	marg. eff..	marg. eff.
Outcome probability	0.048	0.256	0.367	0.329
Male	-0.025***	-0.022	0.038	0.043*
Age at immigration	-0.001	-0.001	-0.001	0.001
High-skilled	0,001	0,001	-0.061*	0.053
Turkish	-0.046***	-0.053	0.098*	0.001
West European	-0.053***	-0.008	-0.047	0.108**
East European	-0.031**	-0.079	-0.037	0.147**
Russian	-0.050***	-0.029	0.234***	-0.155***
Ethnic German	0.010	0.079**	0.055	-0.144***
Rest of the World	-	-	-	-
Social ties	0.039***	-0.016	0.017	-0.039
Observations	1,031			
Pseudo R ²	0.1649			
Log Likelihood	-1,193			

p-value: *** = 0.01, ** = 0.05, * = 0.10

Source: own calculations based on GSOEP 1991-2004.

all on the likelihood to settle in a particular region. However, the ethnicity turns out to be important. A high share of Turkish individuals positively affects the probability of incoming immigrants to locate in this region. Accordingly, the installed basis of immigrants from Eastern Europe as well as from Russia positively affect the likelihood of incoming immigrants in their location choice.

However, so far the relationship between socio-economic characteristics of the immigrants and the location choice as well as alternative-specific attributes of the four different destination regions and the associated location choice have been analyzed separately. In the next subsection, I will present results of the mixed logit models where both location-specific and individual characteristics of immigrants are simultaneously taken into account.

Table 7: Marginal effects of the conditional logit model (Models 3 and 4)

Models	Model 3	Model 4
	marg. eff.	marg. eff.
North Germany	0.199***	0.947***
South Germany	0.241***	0.967***
West Germany	0.237***	0.955***
East Germany	-	-
Unemployment rate	-	-0.114
GDP per employee	-	0.430**
Crimes	-	-0.146**
Share of foreigners	-	-0.312***
Turkish	-	0.759*
West European	-	0.181
East European	-	0.359***
Russian	-	0.989***
Rest of the World	-	-0.228*
Observations	4,124	4,124
Pseudo R ²	0.1122	0.1347
Log Likelihood	-1,268	-1,236

p-value: *** = 0.01, ** = 0.05, * = 0.10

Source: own calculations based on GSOEP 1991-2004.

Mixed logit models

In Table 8 results from two simple mixed logit models are presented. The only difference between model 5 and 6 is the consideration of country of origin effects of the incoming immigrants in the latter model.

In Model 5 immigrant characteristics and attributes of the destination regions are incorporated. For this purpose, I generated interaction terms between individual characteristics and the location factors. In line with Models 3 and 4 the marginal effects for the German regions turn out to be significant showing the expected positive sign since East Germany

Table 8: Marginal effects of the simple mixed logit model (Models 5 and 6)

Models	Model 5	Model 6
	marg. eff.	marg. eff.
North Germany	0.883**	0.493*
South Germany	0.926*	0.605*
West Germany	0.918**	0.529*
North X High-skilled	0.017	0.003
South X High-skilled	0.071	0.021
West X High-skilled	0.044	0.013
North X ethnic German	-0.023	-0.033
South X ethnic German	0.009	-0.038
West X ethnic German	-0.283***	-0.127**
North X social ties	-0.195**	-0.084**
South X social ties	-0.169**	-0.063*
West X social ties	-0.222***	-0.109***
GDP per employee	0.362**	0.137**
Unemployment rate	-0.792	0.081
Crime rate	-0.132***	-0.487**
Share of foreigners	-0.306***	-0.137***
Turkish	0.567	0.532**
East European	0.356***	0.109***
West European	0.168	0.115
Russian	0.907***	0.356***
Rest of the World	-0.187	-0.644
Destination Region X Country of Origin	-	yes ^Δ
Observations	4,124	4,124
Pseudo R ²	0.1606	0.1841
Log Likelihood	-1,199	-1,166

*** = 0.01, ** = 0.05, * = 0.10; Δ = see Table 2.9

Source: own calculations based on GSOEP 1991-2005.

Table 9: Marginal effects of the interaction terms between destination region and country of origin (Model 6)

Model 6	Interaction terms
	marg. eff.
North X Turkish	0.066
South X Turkish	0.084**
West X Turkish	0.068*
North X East European	0.064**
South X East European	0.069**
West X East European	0.081**
North X West European	0.033
South X West European	0.049
West X West European	0.065**
North X Russian	0.067**
South X Russian	0.093***
West X Russian	0.035
Observations	4,124
Pseudo R ²	0.1841
Log likelihood	-1,167

p-value: *** = 0.01, ** = 0.05, * = 0.10

Source: own calculations based on GSOEP 1991-2004.

serves as the reference category. However, interacting the destination regions with the educational background shows that there is no statistical significant difference in the location choice of low- and high-skilled immigrants anymore. Ethnic Germans seem to be indifferent in their location choice with regard to East, North and South Germany. However, they are less likely to move to the region West Germany.

The marginal effects of the variable *social ties* indicates that there exist differences between East Germany and the three other German regions. In East Germany social ties seem to be decisive in the location choice of immigrants. The marginal effect of the unemployment

rate shows that the labor market situation does not seem to be important for the initial location choice whereas the GDP per employee is positive as in the previous models. At first glance, the share of foreigners turns out to be an insignificant explanatory variable. However, distinguishing between the regions/countries of origin I find that the installed basis of East Europeans as well as Russians positively affect incoming immigrants in their location choice.

In Model 6, I additionally control for the country of origin of the arriving immigrants (see Table 9). The results indicate that immigrants from Turkey have a higher probability to move to South Germany than to any other German region. Russian immigrants are more likely to settle in the Southern part of Germany as well. Immigrants from Eastern as well as Western European Countries seem to prefer West Germany as destination region. Overall, the fitting of the model is improved when the country of origin of the incoming immigrants is included (less negative/higher log-likelihood value).

The results of the alternative-specific multinomial probit (ASMNP) model are presented in Table 10. Model 7 represents the counterpart of model 5, i.e. I include the same explanatory variables but I allow for correlation in the error components. In particular, one correlation coefficient is common to all error pairs and the remaining correlations vary over alternatives. Overall, the sign of the marginal effects are similar whereas the size is somewhat different than in the estimation of model 5.

The unemployment rate does not have a statistically significant influence on the location choice probability. However, GDP per employee as well the crime rate show the expected signs. A wealthier region seems to attract incoming immigrants whereas the effect is negative for regions with a high crime rate. With regard to the country of origin of the newly arriving immigrants I find Turkish and Russian individuals to have a higher probability to move to South Germany. For immigrants originating from Eastern European Countries no statistically significant effect is found.

Since model 6 shows a better fitting compared to model 5 it would be straightforward to estimate model 6 with correlation in the error components. However, conducting this estimation I did not obtain any results. In this context, Cameron and Trivedi (2005) underline that the estimation of alternative-specific multinomial probit (ASMNP) models with more

Table 10: Marginal effects of the alternative-specific multinomial probit model with exchangeable correlation (Model 7)

Model 7	East Germany	North Germany	South Germany	West Germany
Correlation of error terms	exchangeable			
	marg. eff.	marg. eff.	marg. eff.	marg. eff.
GDP per employee	-3.485	0.949	1.310	1.235
Unemployment rate	-0.020	0.006	0.008	0.007
Crime rate	2.558**	-0.696**	-0.955**	-0.906**
Share of foreigners	-0.046	0.013	0.017*	0.016*
Turkish	0.013*	-0.035*	0.048**	-0.045*
West European	-0.071	0.019	0.027	0.025
East European	0.076**	-0.021*	-0.028**	-0.026**
Russian	1.705**	-0.465**	-0.636**	-0.604**
Rest of the World	0.101	-0.028	-0.038	-0.036
High-skilled	-0.124			
Ethnic Germans	0.014			
Social ties	0.035***			
Observations	4,124			
Log simulated-likelihood	-1,203			

p-value: *** = 0.01, ** = 0.05, * = 0.10

Source: own calculations based on GSOEP 1991-2004.

than four choices has only had little empirical success because of multidimensional integrals of the likelihood function.

Discussion of results

Overall, the empirical results are quite robust across the different econometric specifications. The estimation of the multinomial logit model where only individual characteristics were taken into account indicates that immigrants from specific source regions, e.g. Russian and Turkish foreign-born, are more likely to move to South Germany. With regard to the educational achievements there only seem to exist small differences in the location choice between low- and high-skilled immigrants. Interestingly, foreign-born who have social ties prior to Germany prior to immigration have a higher probability to move to Eastern Germany. However, overall regions in the Western part of Germany are more attractive to incoming immigrants compared to East Germany as results from the conditional logit model showed. A higher GDP per employee as well as a lower crime rate serve as explanations. However, the unemployment situation seems not to be a decisive variable in the initial location choice of immigrants.

The combined estimation of alternative-specific attributes and individual-specific characteristics allows a better understanding of the initial location choice decisions. In these specifications, both immigrant characteristics as well as location attributes are taken into account. In this context, controlling for the country of origin of the newly arriving immigrants seems to represent an important variable. The results show that immigrants from Turkey are more likely to move to South Germany where a high concentration of Turkish individuals can be observed. Russian immigrants are more likely to settle in the Southern part of Germany for the same reason. The simultaneous inclusion of macro-economic variable such as GDP per employee and the number of crimes also show the expected signs in these econometric models.

However, in order to evaluate the overall fitting of the specifications I compare observed values for the initial location choice with predicted probabilities derived from the estimations (see Table 11). Overall, differences between observed values and predicted probabilities are relatively small. Model 3 where only the locations are included as regressors exactly replicate the observed values for the different location choices.

Table 11: Comparison of model specifications

	East Germany	North Germany	South Germany	West Germany
Observed values	0.052	0.237	0.363	0.348
Model 1 (pred. prob.)	0.053	0.242	0.358	0.347
Log likelihood			-1,273	
Model 2 (pred. prob.)	0.048	0.256	0.367	0.329
Log likelihood			-1,193	
Model 3 (pred. prob.)	0.052	0.237	0.363	0.348
Log likelihood			-1,268	
Model 4 (pred. prob.)	0.04	0.201	0.393	0.302
Log likelihood			-1,236	
Model 5 (pred. prob.)	0.066	0.239	0.373	0.346
Log likelihood			-1,199	
Model 6 (pred. prob.)	0.041	0.245	0.372	0.342
Log likelihood			-1,166	
Model 7 (pred. prob.)	0.045	0.234	0.375	0.346
Log simulated likelihood			-1,203	

Source: own calculations based on GSOEP 1991-2004.

The comparison of the log (simulated-) likelihood shows that model 6 has the least negative value. In this mixed logit model the countries of origin of the immigrants as well as the composition of foreigners in the destination regions are included as explanatory variables. Overall, the simultaneous consideration of location- and individual-specific variables improves the model fitting and thus offers a better understanding which factors are associated with the initial location choice of newly arrived immigrants in Germany.

Model 7 allows for correlations across the utilities in the choice set. However, the relaxation of the independent of irrelevant alternatives assumption does not seem to be decisive in the initial location choice decision of immigrants in Germany. First, the fitting of the model is not improved when estimating the mixed logit model with error components (ASMNP) since the log simulated likelihood is lower compared to model 6. Second, the estimation re-

sults do not significantly change when using the specification with or without correlations in the error terms. Third, the inclusion of additional explanatory variables in the simple mixed logit model (model 6) obviously outweighs its restrictive IIA assumption when compared to the mixed logit model with error components (model 7).

7 Conclusion

The aging and the shrinkage of the population in Germany will come along with shortages in the labor force. Since immigration might mitigate the decline of the regional labor supply I study the initial location choice of newly arrived immigrants for the period 1991-2004. In particular, I analyze the location choice of immigrants in Germany taking individual-specific characteristics as well as location-specific attributes into account.

As a main result, I find that migration networks in Western Germany have a significant impact on the probability of incoming foreign-born originating in the same source country to move to these ethnic clusters. In contrast, in Eastern Germany only concrete social ties between former and newly arrived foreign-born positively influence incoming immigrants to locate in the same region. Referring to the educational background I only find weak evidence that there are differences in the location choice between low- and high-skilled immigrants. In addition, the labour market situation does not represent an important regional attribute for incoming immigrants. However, a high crime rate negatively affects the probability to locate in this region whereas a high GDP per employee is positively associated with the location choice of newly arriving foreign-born.

Although, Germany is exposed to demographic changes immigration is highly regulated since the general recruitment stop for foreign-born established in 1973 is still effective. In addition, with regard to the first and second round of the eastward enlargement of the European Union in May 2004 and in January 2007, respectively, old and new member states agreed on a transition period in the free movement of labor up to seven years. Germany decided to restrict the admission of employees originating from Central and Eastern European Countries (CEEC) to its labor market. However, scarcities in the labor supply which cannot be bridged by natives are likely to result in negative effects on (regional) economic growth.

Hence, the establishment of the free movement of workers between Germany and the new member states of the European Union might alleviate the emerging bottlenecks in the labor force.

References

- ASLUND, O. (2005): “Now and Forever? Initial and Subsequent Location Choices of Immigrants,” *Regional Science and Urban Economics*, 35, 141–165.
- BARTEL, A. (1989): “Where Do the New U.S. Immigrants Live?,” *Journal of Labor Economics*, 7, 371–391.
- BAUER, T., G. EPSTEIN, AND I. GANG (2002): “Herd Effects or Migration Networks? The Location Choice of Mexican Immigrants in the U.S.,” *IZA Discussion Paper 551*.
- BAUER, T., AND K. ZIMMERMANN (1997): “Network Migration of Ethnic Germans,” *International Migration Review*, 31, 143–149.
- BORJAS, G. (1987): “Self-Selection and the Earnings of Immigrants,” *American Economic Review*, 77, 531–555.
- (1995): “Ethnicity, Neighbourhoods, and Human-Capital Externalities,” *American Economic Review*, 85, 365–390.
- CAMERON, A., AND P. TRIVEDI (2005): *Microeconometrics: Methods and Applications*. Cambridge University Press.
- DAMM, A. (2007): “Determinants of Recent Immigrants’ Location Choices: Quasi-Experimental Evidence,” *Journal of Population Economics*, 20, in press.
- DIETZ, B. (1999): “Ethnic German Immigration From Eastern Europe and the Former Soviet Union to Germany: the Effects of Migration Networks,” *IZA Discussion Paper 68*.
- EDIN, F., AND O. ASLUND (2003): “Ethnic Enclaves and the Economic Success of Immigrants - Evidence from a Natural Experiment,” *Quarterly Journal of Economics*, 118, 329–357.

- GREENE, W. (2005): *Econometric Analysis*. Prentice Hall, 5th Edition, Upper Saddle River (NJ).
- GREENE, W. (2008): “Discrete Choice Modeling,” in *The Handbook of Econometrics Vol. 2, Applied Econometrics, Part 4.2*. T. Mills and K. Patterson.
- GRUNDIG, B., C. POHL, AND M. THUM (2007): “Demographic Change and Regional Labour Markets: The Case of East Germany,” *mimeo Dresden*.
- HAIKEN-DENEW, J., AND J. FRICK (2005): *Desktop Compendium to the German socio-Economic Panel (SOEP)*. DIW, Berlin.
- HAUG, S., AND L. SAUER (2006): “Zuwanderung und räumliche Verteilung von Aussiedlern und Spätaussiedlern in Deutschland,” *Zeitschrift für Bevölkerungswissenschaft*, 3.
- KAUSHAL, N. (2005): “New Immigrant’s Location Choices: Magnets without Welfare,” *Journal of Labor Economics*, 23, 59–80.
- LONG, J., AND J. FREESE (2006): *Regression Models for Categorical Dependent Variables using Stata*. STATA Press.
- MCFADDEN, D. (1974): “Conditional Logit Analysis of Qualitative Choice Behavior,” in *Frontiers in Econometrics*. P. Zarembka.
- (2001): “Economic Choices,” *American Economic Review*, 91, 351–378.
- MUNSHI, K. (2003): “Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market,” *Quarterly Journal of Economics*, 118, 549–597.
- NECHYBA, T. U. R. S. (1998): “Community Choice and Local Public Services: A Discrete Choice Approach,” *Regional Science and Urban Economics*, 28, 51–73.
- RYTINA, N. (2006): “Estimates of the Legal Permanent Resident Population and Population Eligible to Naturalize in 2004,” *Population Estimates*, pp. 1–6.
- SCOTT, D., P. COOMES, AND A. IZYUMOV (2005): “The Location Choice of Employment-Based Immigrants Among U.S. Metro Areas,” *Journal of Regional Science*, 45, 113–145.

SJAASTAD, L. (1962): “The Costs and Benefits of Human Migration,” *Journal of Political Economy*, 70, 80–93.

TRAIN, K. (2003): *Discrete Choice Models with Simulation*. Cambridge University Press.

WINTERS, P., A. DE JANVRY, AND E. SADOULET (2001): “Family and Community Networks in Mexico-U.S. Migration,” *Journal of Human Resources*, 36, 159–184.

ZAVODNY, M. (1999): “Determinants of Recent Immigrants’ Locational Choices,” *International Migration Review*, 31, 1014–1030.