

Nowotny, Klaus; Pennerstorfer, Dieter

Working Paper

Ethnic networks and the location choice of migrants in Europe

Working Papers in Economics and Finance, No. 2012-07

Provided in Cooperation with:

Department of Social Sciences and Economics, University of Salzburg

Suggested Citation: Nowotny, Klaus; Pennerstorfer, Dieter (2012) : Ethnic networks and the location choice of migrants in Europe, Working Papers in Economics and Finance, No. 2012-07, University of Salzburg, Department of Social Sciences and Economics, Salzburg

This Version is available at:

<https://hdl.handle.net/10419/71867>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

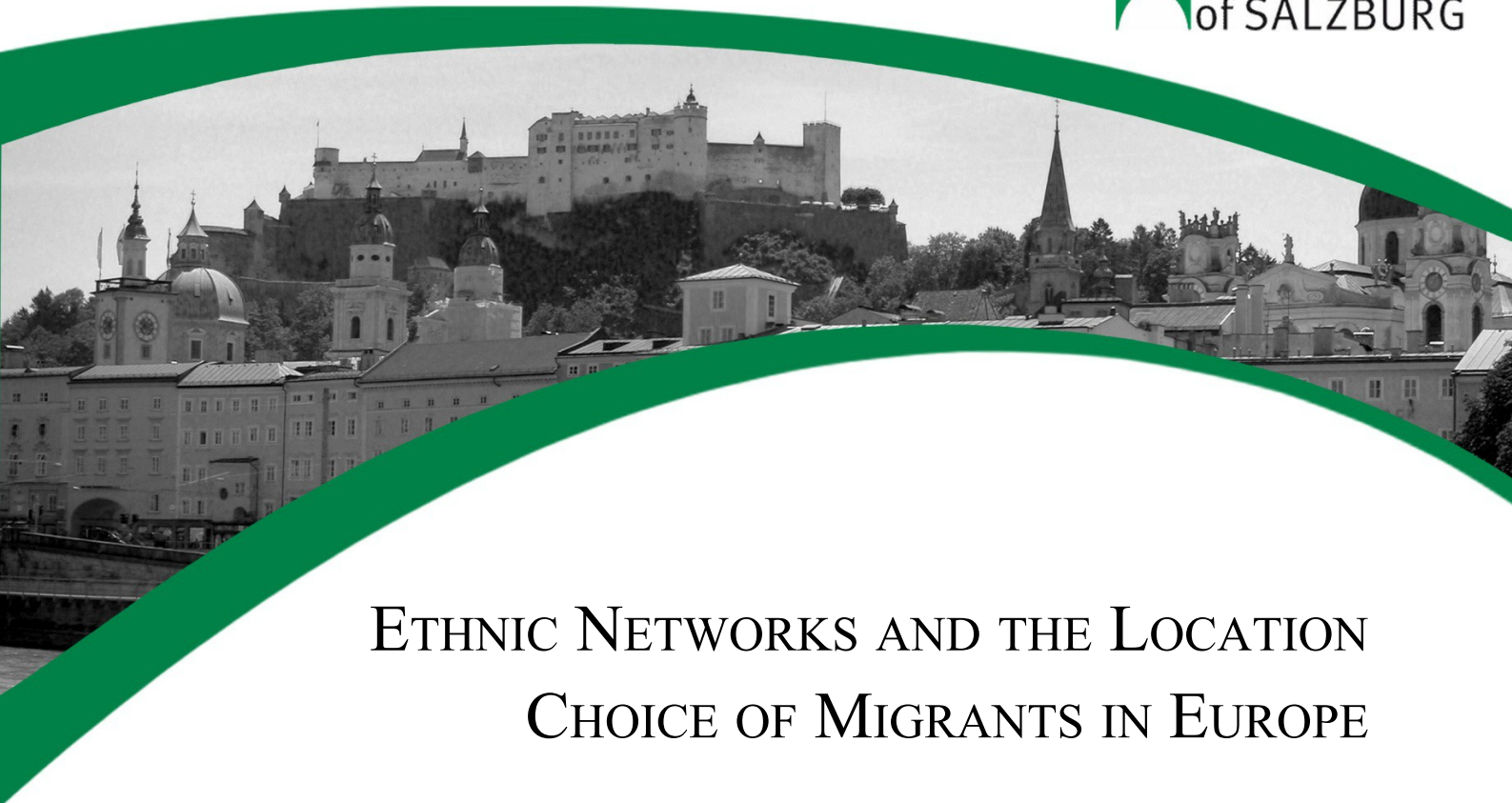
Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



ETHNIC NETWORKS AND THE LOCATION CHOICE OF MIGRANTS IN EUROPE

KLAUS NOWOTNY AND DIETER PENNERSTORFER

WORKING PAPER No. 2012-07

WORKING PAPERS IN
ECONOMICS AND FINANCE

Ethnic Networks and the Location Choice of Migrants in Europe*

Klaus Nowotny[†] Dieter Pennerstorfer[‡]

This version: September 3, 2012

Abstract

This paper analyzes the role of ethnic networks in the location decision of migrants to the EU at the regional level. Using a random parameters logit specification we find a substantially positive effect of ethnic networks on the location decision of migrants. Furthermore, we find evidence of spatial spillovers in the effect of ethnic networks. Analyzing the trade-off between potential income and network size, we find that migrants would require a sizable compensation for living in a region with a smaller ethnic network, especially when considering regions where only few previous migrants from the same country of origin are located.

JEL classification numbers: F22, R23, C35

Keywords: network migration, ethnic networks, random parameters logit

*The authors would like to thank Peter Huber and Jesus Crespo-Cuaresma, participants of the 49th congress of the European Regional Science Association, the 26th Congress of the European Economic Association/65th European Meeting of the Econometric Association and the 2010 Winterseminar of the Gesellschaft für Regionalforschung as well as seminar participants at the University of Innsbruck, the University of Salzburg, the Vienna University of Economics and Business Administration as well as anonymous referees for helpful comments. All remaining errors are our own.

[†]Corresponding author. University of Salzburg and Austrian Institute of Economic Research WIFO. Address: University of Salzburg, Residenzplatz 9, A-5010 Salzburg, Austria. E-mail: klaus.nowotny@sbg.ac.at.

[‡]Austrian Institute of Economic Research WIFO, Arsenal Objekt 20, A-1030 Vienna, Austria. E-mail: dieter.pennerstorfer@wifo.ac.at.

1 INTRODUCTION

Previous research has shown that migrants' choice of location within a country can be explained by differences in economic opportunities across regions (like higher wages or an increased probability of finding employment, see Davies et al., 2001) or by some regions or cities acting as "ports of entry" into a country because of infrastructure endowments (like sea- or airports) or administrative institutions (like central immigration offices). But these factors cannot fully explain the observation that migrants tend to settle where other migrants of the same ethnicity migrated before, resulting in a geographic concentration of migrants with similar ethnicity in specific locations. Since a seminal study on ethnic migrant concentration in the U.S. by Bartel (1989), several studies have formulated hypotheses explaining migrant concentrations theoretically (see Massy et al., 1993, for an overview of some earlier work, or Carrington et al., 1996; Gross and Schmitt, 2003; and Chiswick and Miller, 2005) and have identified the importance of ethnic networks for the location decision of migrants (see, among others, Zavodny, 1999; Bauer et al., 2000; Gross and Schmitt, 2003; Åslund, 2005; Pedersen et al., 2008; Damm, 2009a). Other studies have highlighted the role of ethnic networks for employment and earnings opportunities or educational attainment (see Cutler and Glaeser, 1997; Munshi, 2003; Cardak and McDonald, 2004; Chiswick and Miller, 2005; Damm, 2009b, to name just a few) or the role of "ethnic capital" in determining economic performance (see Borjas, 1992, 1995).

However, the previous literature focused only on local networks and has not considered that the spatial structure of networks around a region may also affect the location decision of newly arriving migrants. But the positive effect of a network is not necessarily limited to regional borders: newly arriving migrants can also benefit from networks in neighboring regions by gaining information on labor market opportunities in these neighboring regions, or by the provision of ethnic goods produced in other regions. Furthermore, some ethnic goods might be provided only if the network size in all regions of a country is large enough.

We thus contribute to the existing literature on the role of ethnic networks on the location decision of migrants by considering not only the size of the local ethnic network as a determinant of migrants' location choice, but also the size of the ethnic network in neighboring regions and other regions of the host country. This is, to the authors' best knowledge, the first article to date which explicitly incorporates this form of spatial heterogeneity in the effect of ethnic networks on migrant's location choice.

Second, we contribute to the existing literature by analyzing location choice at the regional level for European countries, while other studies deal with this

topic either at the national level (Pedersen et al., 2008; Geis et al., 2008) or focus on the regions of a single country (Bartel, 1989; Åslund, 2005; Damm, 2009b). Third, we examine the trade-off between income and ethnic network size and provide an approximation of the compensating variation for changes in the ethnic network size.

As in previous empirical studies (e.g., Davies et al., 2001; Christiadi and Cushing, 2008) the location decisions are estimated at the individual level using a discrete choice model based on random utilities. If our hypothesis is true and networks in neighboring regions matter for the location decision, the independence from irrelevant alternatives property is violated, and the commonly used conditional logit model (McFadden, 1974) is no longer applicable. We therefore follow Gottlieb and Joseph (2006) and apply the more suitable random parameters (mixed) logit framework (see McFadden and Train, 2000).

Based on 2007 data from the European Labour Force Survey our empirical analysis shows that the probability of moving to a region depends not only on the size of the local ethnic network but also—albeit to a smaller extent—on the network size in adjacent regions and other regions of the country. Ignoring the effects of networks in neighboring regions overestimates the effect of network size in the host region and leads to biased results. Deriving the trade-off between income and ethnic network size, we are able to calculate the Euro value of a variation in ethnic networks. Our results show that ethnic networks are highly important for the location decision, and that migrants would require a sizable compensation for moving to a region with a smaller ethnic network.

2 LITERATURE REVIEW

A variety of theoretical and empirical studies has shown that ethnic networks are an important factor in explaining why migrants are concentrated in specific regions of the host country.¹ One of the most frequently cited theories for the importance of networks is that migrant networks produce positive externalities for members of the same ethnic group, so that the costs of migration decrease with the number of previous migrants: networks can provide help with the settlement process, decrease the perceived alienation in the host country (Bauer et al., 2000) or provide financial assistance (Munshi, 2003). Furthermore, networks can provide their members with ethnic goods like food, clothing, social organizations, religious services, media (radio, newspapers, etc.) or marriage markets (Chiswick and Miller, 2005), and the provision of these ethnic goods can be expected to increase with the stock of migrants with similar ethnic background. This creates an externality which provides incentives for other immigrants to settle in regions where they can enjoy a larger supply of ethnic goods.

If there are economies of scale in the production of ethnic goods (as can, for example, be expected for religious services or media), geographic concentration facilitates the supply of these goods at lower prices and reduces the costs of living (especially if ethnic goods make up a large part of the consumption basket), which attracts more immigrants to move into this region even if they could earn a higher wage somewhere else (Chiswick and Miller, 2005).

Ethnic networks also provide information externalities: by being in contact with previous migrants, newly arrived migrants can benefit from a better availability of information on employment opportunities which increases their labor market prospects (Gross and Schmitt, 2003). They can also benefit from job referrals by more established members of the network (Munshi, 2003). Furthermore, if employers with migration background prefer to employ other migrants of similar ethnic origin instead of natives (Andersson and Wadensjö, 2009), a separate migrant labor market can emerge which can even sustain a higher wage than the larger “general” labor market (Gross and Schmitt, 2003).²

A variety of empirical studies in the literature support the network migration hypothesis and find positive effects of ethnic networks on the location decision of newly arrived migrants. However, most of the previous work focuses on the U.S., while there are only few studies covering European countries. Two notable exceptions are Pedersen et al. (2008), who estimate the determinants for migration flows to 22 OECD countries and find a robust and sizable effect of ethnic networks on migration flows, and Geis et al. (2008), who found networks to have a positive (but decreasing in network size) effect on migrant’s choice between four OECD countries (France, Germany, United Kingdom, and the U.S.). Other studies on European countries take a single-country perspective: focusing on Denmark, Damm (2009a) showed that the relocation hazard of refugees randomly assigned to a municipality during the Danish spatial dispersal policy is lower for those assigned to a municipality with a higher percentage of co-nationals, while Åslund (2005) found similar effects for immigrants to Sweden subject to the “Whole of Sweden Strategy” as well as a preference of migrants for regions with larger ethnic networks before the implementation of the strategy.

While there is strong evidence that ethnic migrant networks have a positive effect on the location decision, there can also be negative effects on the utility of both previous migrants (Heitmueller, 2006) and prospective new migrants: continuing migration reduces the income differentials between sending and receiving countries and the wages of migrant cohorts. A similar effect will arise if housing prices increase due to increased demand following an influx of migrants into a region or if the amenities connected to the ethnic network are capitalized into rents and house prices (Gonzalez, 1998). This negative effect of decreasing wages and/or increasing housing prices will at some point dominate the pos-

itive network externality effect, leading to a decline in the attractiveness of a formerly popular ethnic cluster (Portnov, 1999).³ If prospective emigrants take this into consideration when deciding where to locate, an inversely U-shaped effect of network size on the probability of moving to a specific region can arise (Bauer et al., 2007).

3 DATA AND ECONOMETRIC FRAMEWORK

To test for the importance of ethnic networks in the location decision of migrants in Europe we use data from a special evaluation of the 2007 European Labour Force Survey (EU-LFS) which provides not only information on the region of residence (at the NUTS-2 level), but also information on the country of birth. We consider all individuals born outside their country of residence to be “migrants”. The data also allow us to differentiate between those who moved to the EU between 1998 and 2007 and those who have been living in their host country for more than 10 years. The location choice will be modeled for migrants who moved to the EU between 1998 and 2007.

The set of host countries includes only 13 of the 15 EU member states as of 1998⁴ due to missing information on the country of birth in the German and Irish LFS. Using the weight attached to the individual observations in the EU-LFS, we model the location decision of 8,988,710 migrants age 15 or older from 154 countries.⁵

To model the decision of migrant k to settle in region $r \in R$, a random utility framework can be applied (Marschak, 1960) where each of the 158 NUTS-2 regions in the set of alternative regions R yields a region-specific utility U_{kr} .⁶ We impose the simple behavioral model of a utility-maximizing decision maker: migrant k chooses alternative $s \in R$ if and only if $U_{ks} > U_{kr} \forall s \neq r$. Because the decision maker’s utility is not known, observable characteristics of the alternatives X_{kr} can be used to define the representative utility $V_{kr} = V(X_{kr}) \forall r$ which is a function of variables X_{kr} specific to migrant k and region r . Assuming that representative utility is linear in the attributes of the regions, the utility function is given by

$$(1) \quad U_{kr} = V_{kr} + \varepsilon_{kr} = \beta' X_{kr} + \varepsilon_{kr}$$

where ε_{kr} is a random error term.

Ethnic networks

Our main variable of interest in X_{kr} is the size of the ethnic network. We use country of birth to define ethnicity in the EU-LFS data and assume that all individuals born in the same source country belong to the same ethnic group. Although this is only an imperfect measure, it has been widely applied in the literature as a proxy for ethnicity and/or cultural similarity (see, for example, Borjas, 1995; Gross and Schmitt, 2003; Åslund, 2005; Cutler et al., 2008; Damm, 2009a,b, to name just a few).

Because the positive effects of ethnic networks (see section 2) can be assumed to increase with the number of individuals of the same ethnicity living in a region, the most obvious measure of the size of the ethnic network is the stock (the absolute number) of previous migrants with the same country of birth. Since we analyze the location decisions of those who migrated to the EU between 1998 and 2007, the number of previous migrants is proxied by the stock of migrants from the same country of birth who moved to the region before 1998.

Unfortunately, the EU-LFS does not allow a more detailed differentiation by year of arrival in the host country. The only information available in the data is whether migrants who were living in a particular EU country in 2007 migrated before 1998 or between 1998 and 2007. Despite this shortcoming of the data, three arguments justify our approach: first, it takes some time for networks to be effective; only after they have learned the administrative and social conventions of their host country, after they have found jobs or founded businesses providing ethnic goods, etc., previous migrants will be able to provide assistance to newly arrived members of their ethnic community. Second, by including only those who have been living in a region for at least 10 years our network variable includes only the most established members of a migrant's community. Although it could be argued that the tightness of links to the ethnic community decreases over time (for example, if previous migrants assimilate into the host country culture), these established members are likely to be the most helpful for newly arrived migrants. Third, because the network variable includes only those who migrated before 1998, the network size is not affected by those who migrated between 1998 and 2007 for which we model the location decision.

But using the stock of previous migrants from the same country of birth ignores the enormous heterogeneity between the size of ethnic networks in Europe: for example, while (according to the EU-LFS figures for 2007) around 1.17 Mio. Moroccans and around 1.13 Mio. Algerians have been living in the 13 countries considered for at least 10 years (most of them in France), the median number of migrants from the same country of birth who have been living in the 13 countries considered in 2007 is only around 21,000. Therefore, a regional

network consisting of (for example) 10,000 previous migrants may be very important for new migrants coming from a small ethnic group, while it may be relatively irrelevant for someone coming from a very large ethnic group.

Because it is likely that this heterogeneity and the presence of very large ethnic networks affects our empirical results, we normalize the number of previous migrants by dividing it by the total number of previous migrants from the same ethnic group. Although it can be argued that positive network effects depend on the absolute size of the ethnic network rather than on the relative size, the basic effects accredited to the absolute network size also apply to the relative network size: the larger the network (in absolute or relative terms), the more positive externalities are produced for newly arrived migrants. For example, if the provision of ethnic goods depends on the absolute network size, a larger absolute network will produce more ethnic goods compared to other regions with smaller network. But the same applies to a larger relative network: the larger the relative network size in a region, the more ethnic goods will be produced in this region compared to other regions with smaller networks. In both cases, the larger (absolute and relative) network size will make the region more attractive. Furthermore, from the view of a single migrant (which is the focus of our analysis) it is irrelevant whether she compares the absolute or relative network sizes across regions, because she only considers her own ethnic network and the network size across all regions is normalized by the same value. We therefore prefer the relative network size over the absolute network size, but will nevertheless use both definitions in the empirical analysis.⁷

To sum up, for a migrant of ethnic group j , the (relative and absolute) size of the ethnic network size in a specific region s is defined as

$$\text{Relative network}_{js} = \frac{m_{js}^{10+}}{M_j^{10+}} \cdot 100 \quad \text{or} \quad \text{Absolute network}_{js} = m_{js}^{10+}$$

where m_{js}^{10+} is the number of migrants of ethnic group j who have been living in region s for more than 10 years and M_j^{10+} is the total number of migrants of ethnic group j in all of the 13 EU countries considered. Because the marginal utility of networks can decrease with network size (see section 2) the squared network size is also included in X_{kr} .

As outlined in the introduction, the positive effect of ethnic networks does not necessarily end at the region's border. E. g., ethnic goods can also be consumed by individuals living in neighboring regions, or migrants could live in one region and commute to work in a neighboring region where the ethnic network will help them find employment. If there are spatial spillovers, the individual not only considers the size of the local ethnic network, but also the size of the

ethnic network in neighboring regions. We therefore include the sum of the networks in neighboring regions as an additional variable in the regression:

$$\text{Relative network}_{js}^{N_1} = \frac{\sum_{n_1^s=1}^{N_1^s} m_{jn_1^s}^{10+}}{M_j^{10+}} \cdot 100 \quad \text{or} \quad \text{Absolute network}_{js}^{N_1} = \sum_{n_1^s=1}^{N_1^s} m_{jn_1^s}^{10+}$$

with $N_1^s \subset R$ the set of regions sharing a border with region s . Furthermore, we also include the sum of the networks in second neighbor regions N_2^s ($N_2^s \subset R$, $\{s, N_1^s\} \cap N_2^s = \emptyset$):

$$\text{Relative network}_{js}^{N_2} = \frac{\sum_{n_2^s=1}^{N_2^s} m_{jn_2^s}^{10+}}{M_j^{10+}} \cdot 100 \quad \text{or} \quad \text{Absolute network}_{js}^{N_2} = \sum_{n_2^s=1}^{N_2^s} m_{jn_2^s}^{10+}$$

To the authors' best knowledge this is the first article to date which explicitly incorporates this form of spatial heterogeneity⁸ in the context of migrant's location choice.⁹

If there are ethnic goods with strong economies of scale in production the size of the ethnic network in the rest of the host country can also affect the location decision. We therefore also include the sum of the ethnic networks in the rest of the host country ($N_C^s \subset R$, $\{s, N_1^s, N_2^s\} \cap N_C^s = \emptyset$):

$$\text{Relative network}_{js}^{N_C} = \frac{\sum_{l_C^s=1}^{N_C^s} m_{jl_C^s}^{10+}}{M_j^{10+}} \cdot 100 \quad \text{or} \quad \text{Absolute network}_{js}^{N_C} = \sum_{l_C^s=1}^{N_C^s} m_{jl_C^s}^{10+}$$

For networks in neighboring regions we differentiate between neighboring regions in the same country and neighboring regions in other countries. Neighboring countries in other countries are included in the alternative sets $N_1'^s$ (first neighbors) and $N_2'^s$ (second neighbors).

A priori it can be expected that networks in neighboring regions of other countries affect the location choice of migrants to a lesser extent, if at all. Migrants of the same ethnicity living in neighboring countries will not be able to help with immigration issues and bureaucratic structures because of national differences in migration regimes and procedures. Furthermore, labor and housing markets in different countries are subject to different laws, making positive network externalities rather unlikely. National borders will, however, play a lesser role for the consumption of ethnic goods because there are no restrictions on trade and cross-border mobility among EU countries. If significant, the coefficients can, in comparison with their within-country counterparts, provide information about border effects in network externalities.

[Table 1 about here.]

Summary statistics of network size (in 1,000 migrants) are displayed in table 1. The average relative network size is 6.7 % of all migrants of the same ethnicity¹⁰ and the average absolute network consists of about 14,300 individuals. Algerians in the Île de France region (which includes the French capital, Paris) constitute the largest absolute ethnic network: 266,000 Algerians had been living in the Île de France for at least 10 years in 2007. Across all ethnic groups, an average of 17,000 individuals from the same country of birth are living in surrounding regions.

Other explanatory variables

Our choice of other explanatory variables included in X_{kr} follows other studies on this topic (see Bartel, 1989; Grogger and Hanson, 2011; or Davies et al., 2001). Two types of variables will be added to the regression: variables which are specific to the region or country of residence, as well as variables which are specific to a given pair of sending and receiving countries. As will become obvious from the discussion in section 3, variables specific to the sending countries (including sending country fixed effects) cannot be considered in the regressions because they do not vary over alternatives. The same holds true for individual characteristics like age or gender.

Among the region specific X_{kr} attributes is the area of a region (measured in 1,000 km²): if there is a completely even distribution of migrants across space, larger regions will have more migrants than smaller regions. We also control for the population (in 100,000): if urban regions are more attractive, regions with a larger population should (after controlling for region size) attract a higher share of migrants. To control for differences in economic opportunities, we include the unemployment rate (in percent) as well as the average annual income per employee (in € 1,000). Data on population and unemployment (in 2006) as well as average annual income (in 2004) are taken from Eurostat. Regional unemployment rates ranged from 2.3 to 20.2 % in 2006 with an average unemployment rate of 7.3 % (see table 2). The average annual income per employed person was €27,300, and ranged from €10,600 (“Centro” region, Portugal) to €100,000 (Inner London, UK). We expect a negative effect of the unemployment rate and a positive effect of average annual income on the probability of choosing a specific region. We also include a dummy variable for regions in which the capital of a country is located. Capitals can be expected to receive a *ceteris paribus* higher share of migrants because they are the cultural, political and administrative centers of the respective countries.

To control for national differences in laws regarding immigration, labor market access as well as other receiving country-fixed effects, dummies for the receiving countries are included (see also Davies et al., 2001).

[Table 2 about here.]

Among the country-pair specific X_{kr} attributes we include a dummy variable for linguistic closeness taken from CEPII which measures whether a migrant's home and host country share an official language ($= 1$, zero otherwise). A common language not only reduces the costs of migration considerably (see Pedersen et al., 2008), but it can also raise the returns-to-skill in the host country (Grogger and Hanson, 2011). We also include a neighborhood dummy which is 1 if the host and home countries share a common border, and zero otherwise. Again, a positive effect can be expected because a common border facilitates not only legal, but also illegal immigration and can thus lead to *ceteris paribus* higher migration. (Former) colonial ties between two countries can also affect the location choice of migrants. Data on colonial relationships are also taken from CEPII and we include a dummy variable capturing whether two countries were in a colonial relationship after 1945 ($= 1$, zero otherwise). To proxy for the costs of migration (or the costs of visiting relatives at home), the distance (in 1,000 km, as the crow flies) between the capital of the sending country and the geographical center of the region of residence is also included, as is the squared distance. We expect a negative (but possibly decreasing) effect of distance.

Representative utility V_{kr} is thus a linear function of receiving region specific variables (ethnic networks, area, population, average income, unemployment, capital city), receiving country specific variables (country dummies) as well as country-pair specific variables (common official language, common border, colonial ties after 1945, distance) which are assumed to determine the location choice of migrants.

Econometric method

Assuming that the random utility term ε_{kr} in equation (1) is i.i.d. extreme value, the probability that individual k chooses location s , P_{ks} , could be estimated by a conditional logit (CL) model (McFadden, 1974):¹¹

$$(2) \quad P_{ks} = \frac{\exp(\beta' X_{ks})}{\sum_{r=1}^R \exp(\beta' X_{kr})}$$

One feature of this approach is that all variables which do not vary across alternatives (such as individual or sending country characteristics) are cancelled out in the CL formula (2). This not only allows the estimation without source

country data, which seriously reduces the amount of data required and allows the analysis of location choice based on receiving region characteristics alone (cf. Ortega and Peri, 2009), but—more importantly—also controls for any unobservable individual or sending country characteristics which could otherwise lead to omitted variable bias in a cross-sectional setting.

As is well known, in the conditional logit model the odds ratio between two alternatives s and t depends only on the characteristics of s and t and not on the availability or characteristics of other alternatives, a property known as “independence from irrelevant alternatives” (IIA):

$$(3) \quad \frac{P_{ks}}{P_{kt}} = \frac{\exp(\beta' X_{ks}) / \sum_{r=1}^R \exp(\beta' X_{kr})}{\exp(\beta' X_{kt}) / \sum_{r=1}^R \exp(\beta' X_{kr})} = \frac{\exp(\beta' X_{ks})}{\exp(\beta' X_{kt})}$$

However, if ethnic networks in neighboring regions matter, the probability of choosing region s no longer depends on the characteristics of s alone, but also on the characteristics of the neighboring regions $R(s) = \{N_1^s, N_2^s, N_C^s, N_1'^s, N_2'^s\}$. Similarly, the probability of choosing region t depends on the attributes of t 's neighbors $R(t) = \{N_1^t, N_2^t, N_C^t, N_1'^t, N_2'^t\}$. The odds between s and t are then given by:

$$\frac{P_{ks}}{P_{kt}} = \frac{\exp(\beta_1' X_{ks} + \beta_2' X_{kR(s)})}{\exp(\beta_1' X_{kt} + \beta_2' X_{kR(t)})}$$

which violates the IIA property: the ratio of the probabilities no longer depends on the characteristics of s and t alone, but also on the characteristics of the regions in $R(s)$ and $R(t)$.¹²

This calls for a model which does not exhibit the IIA property. Probably the most flexible model is the random parameters logit (RPL, see McFadden and Train, 2000; Hensher and Greene, 2003; Train, 2009, for an overview).¹³ Although the random parameters logit framework goes back to the early 1980's (among the first applications are Boyd and Mellman, 1980, and Cardell and Dunbar, 1980) and recent advances in simulation techniques (foremost, the use of Halton draws, see below) and computing power have made its estimation more practicable, applications of the random parameters logit model are still scarce in migration research (one notable exception is the paper by Gottlieb and Joseph, 2006).

The random parameters model can be derived from utility-maximizing behavior by allowing the parameters of the characteristics X_{kr} in the representative utility function to vary over individuals.¹⁴

$$U_{kr} = \beta_k' X_{kr} + \varepsilon_{kr}$$

In this utility function, β_k is a vector of coefficients for individual k representing k 's preferences. The utility function is thus heterogeneous across individuals, and the coefficient of a regional characteristic can not only have a different magnitude for different individuals, but also a different sign. The coefficients in β_k are assumed to vary over decision makers with density $f(\beta|\theta)$, where θ are the parameters describing the density of β . As in the conditional logit model, ε_{kr} is assumed to be i.i.d. and follow an extreme value distribution. If the β_k 's were known, the probability of choosing a specific region s would be given by:

$$(4) \quad L_{ks}(\beta_k) = \frac{\exp(\beta'_k X_{ks})}{\sum_{r=1}^R \exp(\beta'_k X_{kr})}$$

As is obvious from equation (4), the RPL model shares the advantages of the CL model: all variables which do not vary across alternatives are cancelled out which controls for both observable and unobservable individual and source country characteristics even in a cross-section. In addition, it does not impose the IIA property (which would lead to inconsistent estimates if violated) and thereby allows for unrestricted substitution patterns between alternatives, and it allows for heterogeneity in the effects of explanatory variables on the probability of choosing a specific region.

Because the β_k 's are unobserved the probability of choosing s is the integral of (4) over all possible values of β_k (Train, 2009, p. 138):

$$(5) \quad P_{ks} = \int \left(\frac{\exp(\beta'_k X_{ks})}{\sum_{r=1}^R \exp(\beta'_k X_{kr})} \right) f(\beta|\theta) d\beta$$

Because the integral in (5) does not have a closed form solution, it must be approximated through simulation. Simulation is based on drawing a value of β from $f(\beta|\theta)$ and using this draw to calculate the logit probability in (4). This step is repeated many times, and the average computed value of $L_{ks}(\beta_k)$ gives the simulated probability \check{P}_{ks} which can be inserted into the simulated log likelihood

$$(6) \quad SLL(\theta) = \sum_{k=1}^K \sum_{s=1}^R y_{ks} \ln \check{P}_{ks}$$

The maximum simulated likelihood estimator is the value of θ that maximizes the simulated log likelihood (see Train, 2009, 144) and can be estimated for example in the STATA statistics package using the estimator by Hole (2007).

The mixing distribution $f(\beta|\theta)$ describes the distribution of a parameter β in the population. If the parameters are assumed to be normally distributed, the estimated θ s are the mean and standard deviation of a normal distribution.

In the econometric model we follow Gottlieb and Joseph (2006) by specifying some coefficients as fixed and the rest as normally distributed.¹⁵ A fixed parameter is essentially a normally distributed parameter whose standard deviation is restricted to zero (Hensher and Greene, 2003), and for which only a mean will be estimated. We assume the coefficient of area (in 1,000 km²) to be fixed and the same for all individuals: if migrants were evenly distributed across space, larger regions would have a *ceteris paribus* higher probability of being chosen, independent of individual tastes. The country-specific dummy variables are also treated as being fixed.¹⁶ All other coefficients are unrestricted and assumed to be normally distributed.¹⁷ The estimated parameters θ for these coefficients are thus the mean and standard deviation of a normal distribution. This allows us to calculate the area of the density function $f(\beta|\theta)$ which is below and above zero. As mentioned above, in the random parameters logit a coefficient is not necessarily positive or negative for all individuals. If part of the area of $f(\beta|\theta)$ is below zero, a variable constitutes an attractor for some, and a repellent for other individuals.

Our simulation uses quasi-random Halton sequences (Halton, 1960), which is considered more effective than simulation based on random draws (see Bhat, 2001; Train, 1999, 2009; Hensher, 2001). Although there is no general agreement on the number of Halton draws to be used to achieve stable parameters, Hensher and Greene (2003, p. 154) note that models with a small number of alternatives and random variables can “produce stability with as low as 25” Halton draws per observation, and that “100 appears to be a ‘good’ number”. Train (2009, p. 230) notes that “[...] a researcher can expect to be closer to the expected values of the estimates using 100 Halton draws than 1000 random draws”. However, the number of required draws will be higher the more complex the model (Hensher and Greene, 2003, p. 154), so that these results cannot be generalized. We use 500 Halton draws for the simulation of the random parameters logit model.¹⁸

4 ESTIMATION RESULTS

Relative network size

Table 3 shows the results of the random parameters logit regression estimating the location choice of migrants who moved to the 13 host countries considered between 1998 and 2007 using our preferred measure of relative network size. In addition to the estimated random parameters’ mean and standard deviation (which define the distribution of the coefficients in the population), table 3 shows the proportion of the parameters’ distribution which is above zero. This gives us the percentage of the population for which the parameter is positive.

The fifth column gives the exponentiated means of the random parameters, which can be interpreted as mean odds ratios. Finally, the last two columns give the coefficients and odds ratios of a conditional logit regression. Although the conditional logit’s IIA assumption is violated if our hypothesis of spatial spillovers in network effects is correct, the conditional logit can still serve as an approximation to a model which relaxes this assumption (cf. Dahlberg and Eklöf, 2003). Due to lack of space, the estimated receiving country fixed effects are reported in table A1 in the appendix.¹⁹

[Table 3 about here.]

The results of the random parameters logit support our hypotheses: not only does a larger ethnic network attract more migrants to a region, the estimated probability of choosing a specific region also increases with ethnic networks in neighboring regions. All else equal, at the mean parameter value the odds of choosing a region are 45.6 % larger if the total share of individuals from the same ethnic background in the region increases by 1 percentage point (p.p.). The effect of the ethnic networks is, however, decreasing in network size, as indicated by the negative coefficient of the squared network variable. Furthermore, at the mean parameter value the odds ratio is 5.3 % larger if the ethnic network in neighboring regions increases by 1 p.p., and even a 1 p.p. increase in the ethnic network of second neighbors is still associated with 3.7 % increase in the relative odds at the mean parameter. Networks in the rest of the country also play a role for the location decision, but the effect is rather small. Ethnic networks in neighboring regions of other countries also affect the location decision positively, but the estimated coefficients are smaller than for within-country neighbors, which points to substantial border effects in the influence of ethnic networks.

The positive effects of ethnic networks thus extend beyond regional borders, although the coefficients of networks in neighboring regions are considerably smaller than the coefficient for the network in the region of residence. This does not mean, however, that they should be excluded from the analysis: excluding the spatially lagged network size variables increases the (mean of the) coefficient of the local network variable from 0.376 to 0.651 and reduces the coefficient of the squared network variable from -0.017 to -0.044 .²⁰ This indicates that there would be an omitted variable bias if the spatially lagged network variables are ignored, which would lead to an overestimation of the effect of local ethnic networks.

As mentioned above, the RPL model estimates both the mean and the standard deviation of our random parameters. Despite most estimated standard deviations being rather small, table 3 shows that they are statistically significant, so that there is at least some degree of heterogeneity between individuals.

Nonetheless, ethnic networks are an attractor for all individuals: 100 % of the normal distributions of most network coefficients are above zero. The only exception are the coefficient of the size of the ethnic network in the rest of the country, which is negative for about 23.1 %, and the parameter of network size in second neighbors of another country, which is negative for a small proportion of migrants. In addition, for about 10.2 % of migrants the coefficient of the squared network term is positive, indicating that the utility of these individuals increases exponentially with ethnic network size.

Concerning the other variables, the RPL regression shows that migrants *ceteris paribus* prefer regions with more inhabitants, lower unemployment rates and higher average income, as expected. The effect of regional size is negative: after controlling for population size, migrants prefer smaller regions, which implies that (urban) regions with a higher population density are more attractive for migrants, for example because thick labor markets in urban areas make it easier for them to find employment (Moretti, 2010). The combined effect of population and area size can also account for the result that regions with capital cities exert a positive influence on the location decision of only about 30.0 % of all migrants: *ceteris paribus*, capital cities are not necessarily more attractive for the majority of newly arriving migrants when compared to regions of the same size and population.

As expected, distance—our proxy for the costs of migration—has a negative, but decreasing effect on the location decision. A common official language increases the odds of choosing a specific region, but only for about three quarters (75.4 %) of the migrants. In contrast to results found by previous research (see, for example, Ortega and Peri, 2009; Grogger and Hanson, 2011), a past colonial relationship between the sending and receiving countries is an attractor for only about 11.0 % of the migrants. One explanation for this finding is that a common colonial history is highly correlated with having a common official language, so that most of the effect of a shared history are already captured by the common official language dummy.

Comparing the results of the random parameters to the conditional logit regression shows that the differences between the mean RPL estimates and the conditional logit are quite substantial for some coefficients, especially those of the network size variable and parameters with a high degree of heterogeneity in the population (such as the capital, common border, and colonial relationship dummies). The evidence provided in this paper does thus not lend support to the hypothesis that a CL model can be used as an approximation to the RPL model which relaxes the IIA assumption, but rather shows that imposing a conditional logit (which implies fixed coefficients) on a empirical model characterized by a high degree of heterogeneity in the coefficients can lead to a severe bias. For

example, the odds ratio of the capital dummy variable is considerably smaller in the RPL than in the CL regression. In another example, a common border has a positive effect on the location decision in the CL model, while the mean RPL estimate is negative: only 46.1 % prefer regions in neighboring countries.

As is well known, the ratio of two parameters in a logit model can be used to calculate the trade-off between two variables x_{1kr} and x_{2kr} (see Davies et al., 2001; Train, 2009). Setting the total derivative of the logit probability to zero and solving for the change in x_{1kr} that keeps the probability of choosing region r constant following a change in x_{2kr} yields:

$$(7) \quad \left. \frac{dx_{1kr}}{dx_{2kr}} \right|_{dP_{kr}=0} = -\frac{\beta_{2k}P_{kr}(1-P_{kr})}{\beta_{1k}P_{kr}(1-P_{kr})} = -\frac{\beta_{2k}}{\beta_{1k}}$$

Using a cost or income measure as x_{1kr} this trade-off can be interpreted as the willingness-to-pay or compensating variation (CV, see Dahlberg and Eklöf, 2003; Sillano and Ortúzar, 2005): if x_{1kr} is income, ratio (7) gives the amount of money which would compensate an individual for a increase in x_{2kr} by one unit. If both parameters are positive, the compensation for an increase in x_{2kr} is negative, as expected.

Ratio (7) can be used to calculate the amount of income which would compensate an individual for moving to a region with a smaller ethnic network or, equivalently, the amount of income the individual would be willing to forego in order to live in a region with a larger ethnic network. Because the parameters in the RPL are random variables which vary across the population with density $f(\beta|\theta)$ we follow Sillano and Ortúzar (2005) and calculate the compensating variation from the individual-level parameters derived from the simulation model (see Train, 2009). Migrant k 's compensating variation for a change in the size of the ethnic network can then be calculated as

$$CV_k(\text{Network}_{js}) = -\frac{\gamma_{1k} + 2\gamma_{2k}\text{Network}_{js}}{\mu_k}$$

where γ_{1k} is individual k 's coefficient of the network variable, γ_{2k} her coefficient of the squared network variable and μ_k her coefficient of the average income variable. The compensating variation therefore depends on the size of the network, and will decrease with network size if the squared network parameter is negative (as is the case for about 90 % in our sample, see table 3).

Table 4 shows the compensating variation calculated from individual level parameters at different network sizes as well as the calculated compensating variation for ethnic networks in neighboring regions and the rest of the country. They give the amount of annual income an individual would require as compensation for moving to a region where the ethnic network is 1 p. p. smaller.

[Table 4 about here.]

The calculation based on individual level parameters reveals a sizable compensating variation at small network sizes: the amount of potential earnings which would compensate an individual for a 1 p.p. lower network size is, on average, about € 21,600 at a network size of 1 %. Thus, when choosing between two otherwise equivalent regions where one region has a network size of 1 % and the other a network size of 0 %, the probability of moving to these regions would only be equal to the individual if the expected annual income in the region without an ethnic network is € 21,600 higher than the expected income in the other region. This is only slightly lower than the mean average annual income per employee of € 27,300 (see table 2). While such a large willingness to pay would be unrealistic, this underlines the finding that regions without networks are highly unattractive, and that ethnic networks are so important that regions without an ethnic network must provide considerably better income opportunities to be considered equally attractive.

As the network size increases, the compensating variation decreases considerably, which corresponds to the findings of Gonzalez (1998). At the same time, the standard deviation of the CV estimates increases with network size, which reflects the considerable heterogeneity in the individual squared network parameters. At a network size of 5 % the compensating variation drops to about € 13,000 on average, and to about € 2,300 at a network size of 10 %. At the mean network size of 6.65 % the average compensating variation is about € 9,500. The distribution of the compensating variation at the mean network size is depicted in figure 1.

[Figure 1 about here.]

Table 4 also shows that the compensating variation for (the sum of) the network sizes of neighboring regions is considerably smaller than the CV for the network within a region. Furthermore, the compensating variation decreases with distance, and there is a sizable difference between the CV for networks in neighboring regions compared to regions in the rest of the country. In addition, there is also a border effect: the compensating variation for networks in neighboring regions of the same country is about € 3,200, while it is only € 2,300 for neighboring regions in another country. The same pattern holds for networks in second neighbor regions. These results show that the importance of networks decreases with distance to the region of residence, and that there are sizable border effects in the spatial spillovers of ethnic networks.

Absolute network size

As outlined in section 3, we prefer the relative network definition as measure for the size of an ethnic network. Nevertheless, we also estimated the model using absolute network size (see table 5). The main results of section 4 are unaltered by this change in the definition of the network variable: ethnic networks both in the same region as well as in neighboring regions (both within the country and across borders) significantly affect the location choice of migrants.

[Table 5 about here.]

The odds of choosing a region are 33.2 % larger if the local ethnic network increases by 1,000 individuals, and the effect of network size is again positive for all migrants. This also holds true for the effects of networks in neighboring regions which are, however, of limited importance in this specification: increasing the network size in a neighboring region by 1,000 individuals increases the odds of choosing a region by only 0.2 %. Networks in the rest of the country also play a role for the location decision, the effect is, however, again rather small and even negative for about 63 % of all migrants. Compared to the estimation results of table 3 the proportion of individuals for which this coefficient is positive decreases by about 40 p. p. in reaction to the change in network definition.

Compared to the regression using relative network size (table 3), networks in neighboring regions and the rest of the country are thus of lesser importance (relative to the effect of the network in the region of residence) if absolute network size is used. This could imply that networks in neighboring regions are more important for small ethnic groups, while for members of ethnic groups with a high number of previous migrants (which are likely to dominate the results in table 5) networks in neighboring regions are of lesser importance.

Again, ethnic networks in neighboring regions of other countries affect the location decision positively. In contrast to the relative network size regression (table 3), the estimated coefficients are slightly larger than for within-country neighbors, but the absolute network size estimation shows a higher degree of heterogeneity in these parameters. Again, the empirical model shows evidence in support of the optimal network size hypothesis because of the negative parameter of the squared local network.

A direct comparison of the random parameters and conditional logit regressions again shows that imposing fixed parameters on the empirical model would lead to severely biased coefficients, especially for parameters with a high degree of heterogeneity, such as the capital, common border, and colony dummies, but also in the absolute network size variable.

As before, we can use the parameters estimated to derive a measure for the amount of income that would compensate migrants for a decrease in the eth-

nic network. At the average absolute network size of about 14,300 migrants (see table 1) the compensating variation for a decrease in the ethnic network in the same region of 1,000 individuals is about €7,700. Again, this sizable compensating variation shows that networks are very important for newly arrived migrants, so that they are willing to forego a sizable part of potential income in order to live in a region where the ethnic network is larger.

5 ROBUSTNESS

Alternative definition of neighboring regions

In the previous regressions, the spatial structure of ethnic networks was captured by summing up the network size in neighboring NUTS-2 regions. This ignores that the size of NUTS-2 regions differs across countries. Although the “Nomenclature des Unités Territoriales Statistiques” (NUTS) should ensure at least some comparability across regions in the European Union, NUTS-2 regions across Europe are quite heterogeneous. While continental France is more than 1.5 times the size of Germany, there are 39 German and only 22 (continental) French NUTS-2 regions. In another example, while the Region Övre Norrland (NUTS-2 code: SE33) had an area of about 165,300 km² and about 3.3 inhabitants per km² in 2007, the region Bruxelles-Capitale (NUTS-2 code: BE10) had 161 km² and about 6,500 inhabitants per km² (in 2007) according to Eurostat data. This of course implies that the availability of networks in neighboring regions may differ with region size.

To test whether this is the case we restrict the set of neighboring regions to include only regions whose geographical center is within a radius ρ of 0–100 or 101–200 kilometers (as the crow flies) from the geographical center of the region of residence. The networks in all other regions (outside the 200 kilometer radius) are included as an additional regressor. As shown by the summary statistics in table 6, on average 4.0 % of an ethnic network (outside the region of residence) can be found within a 100 kilometer radius. The average number of those living in regions within 101–200 kilometers is about 5.1 %. As before, we differentiate between ethnic networks within the host country and networks in neighboring countries. Networks in neighboring regions of other countries are considered only when they are within a radius ρ' of 0–100 and 101–200 kilometers. As before, the size of the local ethnic network in the region of residence as well as its squared value are included, so that the regression parameters can be directly compared with the results in table 3.

[Table 6 about here.]

Despite the change in the definition of neighboring regions, the main conclusions are unaltered (see table 7): even if networks in neighboring regions are considered only if they are within a given distance to the region of residence, they still affect location choice positively for all migrants. Networks within a 100 kilometer radius exert a larger influence than networks within a 200 kilometer radius or networks in the rest of the country, so that the effect of networks decreases with distance from the region of residence. Again, the largest effect can be found for ethnic networks in the migrants' region of residence. As before, this effect is decreasing in network size for most migrants, and only for 10 % the squared network variable has a positive coefficient. The estimated network parameter is slightly larger in this regression, but close to the original parameter of table 3.

[Table 7 about here.]

In contrast to the previous regressions, networks in neighboring regions of other countries within a 100 kilometer radius affect location choice negatively for 84 % of all migrants. Networks in regions of other countries within in a 100–200 kilometer radius again exert a positive influence on the probability of choosing a specific region. This can be explained by the differences in coverage between the two definitions: on average, each region shares a border with 0.53 regions in other countries, but only 0.29 regions in other countries are (on average) within a 100 kilometer radius.²¹ Overall, there are only 21 regions where the closest region in another country is less than 100 kilometers away and actually hosts a network. The majority of these regions (16) are in Belgium and the Netherlands, and it is therefore likely that the difference to the main regression arises from the specifics of these countries or the ethnic groups living in these countries.

Migrants and refugees

Because we cannot distinguish between migrants and refugees in the EU-LFS data, our results might partly be driven by the differences in the location decisions of these groups. For example, in some European countries (especially Sweden, Denmark and the Netherlands) spatial dispersion policies are or have been in place which restrict(ed) the freedom of movement of refugees. Therefore, we also test the robustness of our results by estimating the model on subsets of origin countries defined by the level of development. We assume that individuals born in countries with low levels of development are more likely to be refugees who cannot freely choose their residence location. Migrants from these countries are therefore excluded from the regression.

We use the United Nation’s Human Development Index (HDI) to define the sending country’s level of development.²² The HDI ranges between zero and one, and countries are classified into one of the following four categories: very high human development ($0.9 \leq \text{HDI} \leq 1$), high human development ($0.8 \leq \text{HDI} < 0.9$), medium human development ($0.5 \leq \text{HDI} < 0.8$) and low human development ($0 \leq \text{HDI} < 0.5$).

[Table 8 about here.]

To distinguish refugees from migrants, we estimate the model only for individuals born in countries with at least medium human development. The results in table 8 show that excluding those most likely to be refugees increases the size of the estimated parameter for local networks only slightly from 0.376 to 0.427. At the mean estimated parameter value, the odds of moving to a region are 53.3 % higher if the size of the ethnic network increases by 1 p.p. if refugees are excluded, an increase of 7.7 p.p. compared to the regression for all countries (table 3). The other estimated parameters are hardly affected by excluding refugees. Therefore, the main conclusions of our preferred specification remain intact.

Educational and retiree migration

Finally, to focus specifically on the effects of networks on labor migration the model was also estimated on a restricted sample including only migrants between 25 and 54 years of age. Both the location choice of younger migrants as well as the location choice of older migrants may be driven by characteristics not related to the labor market. For example, younger individuals moving abroad to study will choose their location based on education opportunities, and not based on regional labor market characteristics. The location choice of retired individuals, on the other hand, may be driven by factors such as climatic conditions, as evidenced not only by retiree migration to Florida but also by international retiree migration to mediterranean countries in Europe (see Warnes, 2009, for a recent review).²³

[Table 9 about here.]

The regression on the restricted sample shows that the main conclusions of our analysis are practically unchanged (table 9). As before, both local ethnic networks as well as ethnic networks in neighboring regions have a significantly positive effect on the location choice of working-age migrants to Europe. Furthermore, quantitative differences between the estimated parameters are rather small, supporting the results of our main specification (see table 3).²⁴

6 CONCLUSIONS

This paper analyzed the effect of the spatial structure of ethnic networks on the location decision of migrants who moved to the EU between 1998 and 2007. We used a random parameters logit specification which not only allows for heterogeneity in utility functions but also controls for both observable and unobservable individual and sending country characteristics in a cross-sectional setting.

The regressions show a substantially positive effect of ethnic networks on the location decision of migrants, providing strong evidence for ethnic clustering of migrants among European regions. The estimated model also reveals substantial variations in taste across individuals. The effect of ethnic networks in the same region is, however, always positive for all individuals. We also find evidence of spatial spillovers in the effect of ethnic networks: networks in neighboring regions (both in the same country as well as across the border) and networks in the rest of the country significantly help to explain migrants' choice of target regions. The spatial structure of ethnic migration thus matters for the location decision, and the positive effects of ethnic networks extend beyond regional and national borders. Additional estimations using different network and neighborhood definitions as well as regressions on subsamples of the data confirm the robustness of our findings.

The empirical evidence of spatial spillovers violates the IIA assumption, rendering the more conventional conditional logit model invalid. This makes random parameters logit, which is robust to IIA violations, the appropriate model for estimating location choice at the individual level when the characteristics of neighboring regions affect the location decision. Furthermore, the significant standard deviations in the random parameters logit show that the limitations imposed by conditional logit on the individual parameters are too strict. We therefore conclude that the random parameters logit is superior to the conditional logit in the analysis of the location decision of migrants and that there are considerable differences between conditional and random parameters estimates if there is a high degree of heterogeneity in the population and if there are spatial spillovers.

Comparing the effect of networks to the effect of potential income differences, we find that migrants would require a sizable compensating variation for changes in the network size, especially for regions where only few previous migrants from the same country of origin are located. At the mean network size the average compensating variation for a 1 percentage point decrease in the size of the ethnic network is about € 9,500. This shows that ethnic networks play a very important role in the location decision and that regions without a network are highly unattractive. This implies that migrants should be highly

concentrated in specific regions with larger networks. But—as implied by the negative sign of the parameter for squared network size—if networks become too large, the attractiveness of a region will decline, lending support to the inversely U-shaped effect of network size found by Bauer et al. (2007). There is, however, a considerable heterogeneity in the compensating variation across individuals.

While networks in neighboring regions matter for the location decision, the compensating variation for networks in neighboring regions is considerably smaller than for networks in the region of residence. We also find a substantial border effect for the effects of ethnic networks. Our results therefore show that ethnic networks in neighboring regions matter, but that the importance of networks decreases with distance to the region of residence, and that national borders reduce the positive effect of ethnic networks in neighboring regions.

Our results point to a strong “lock-in effect” of the ethnic structure of migration, so that the current ethnic structure in part determines the future regional pattern of migration. This implies that the heterogeneous use of restrictions on the movement of labor among the EU-15 countries during the transitional period can be expected also to have long-run effects on the patterns of migration from the 8 member states which joined the EU in 2004. But regional concentrations of migrants of the same ethnicity can be detrimental to integration measures and foster the evolution of parallel societies. However, spatial dispersion policies which aim at breaking up regional patterns of ethnic migration will lead to a substantial welfare loss for migrants, which must be considered in the evaluation of such policies.

There is, of course, scope for future extensions. First, there may be differences according to education level of migrants. E.g., highly skilled migrants may avoid regions with large concentrations of low-skill migrants of the same ethnicity to escape statistical discrimination (cf. Stark, 1994), so that it would be interesting to analyze the location patterns of migrants of different skill levels. Second, it could be interesting to analyze the substitution patterns between regions based on the random parameters logit model. Analyzing these substitution patterns could, for example, shed light on the effects of changes in economic conditions (or migration policy) in one country (or region) on migration to all other countries (or regions) and thus provide an important tool to forecast future migration patterns based on past migration. Third, although our empirical results clearly show that there are spatial spillovers in the effects of ethnic networks, more research is still needed on how exactly the spatial structure of ethnic networks works and how newly arriving migrants benefit from networks in neighboring regions.

Notes

¹As an alternative to network effects, Bauer et al. (2005) and Epstein (2010) argued that herd behavior can constitute another explanation for the creation of ethnic clusters in specific regions. Herd behavior and network effects are—although conceptually different—not mutually exclusive: both effects can exist simultaneously and determine the location decisions of migrants. The presence of network externalities in this context can even increase the probability that herd behavior will be observed (Epstein, 2010).

²Edin et al. (2003) found empirical support for a positive effects of ethnic networks on migrant earnings. In an analysis of Mexican migrants in the U.S., Munshi (2003) provides evidence that networks not only increase the probability of employment, but also help to channel network members into higher paying occupations. Bartel (1989, p. 388), on the other hand, showed that clustering negatively influences the economic success of migrants. One explanation for this is that migrant clusters are negatively correlated with foreign language fluency (Lazear, 1999), which is in turn a prerequisite for entering the host country’s labor market (see also Bauer et al., 2005; Bisin et al., 2011). Gonzalez (1998) also found a negative effect of enclave size and wages among Mexican migrants in the U.S. Damm (2009b) concludes that the positive effects of ethnic networks more than outweigh the negative effects, and that all things considered living in a region with a larger ethnic network has a positive effect on wages.

³Local ethnic networks can, however, still grow beyond this optimal size if the region still provides a higher utility compared to all other available regions, even if new migrants take into account that their utility will decrease with every other migrant that follows (Heitmueller, 2006).

⁴Austria, Belgium, Denmark, Finland, France, Greece, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden and the United Kingdom.

⁵Serbia, Montenegro and the Kosovo are considered a single source country because of data restrictions. The choice of migrating vs. staying in the home country is not modeled.

⁶Overseas territories as well as the Spanish exclaves Ceuta and Melilla are not included in R . The same holds true for the relatively remote Canary Islands and the Azores and Madeira island regions. Åland (Finland) as well as the Highlands and Islands and North Eastern Scotland regions in the U. K. must be excluded because of lacking data. Furthermore, Denmark must be considered a single NUTS-2 region.

⁷We do not consider the proportion of migrants of the same ethnicity among the population in a region, which could be used as yet an alternative definition. This proportion could be interpreted as an indicator for the number of possible interactions with same-ethnicity individuals in random encounters in the region (see, for example, Lazear, 1999). We believe this measure to be of lesser importance to the decision maker, not only because regional migrant networks tend to be spatially concentrated even within the region (e.g., ethnic enclaves such as the Chinatowns in U.S. cities like San Francisco or New York or in European cities like Liverpool and London are well defined within a few city blocks) which makes random encounters less important, but also because it is unrelated to some of the positive network effects mentioned in the literature (see section 2), for example the provision of ethnic goods. We therefore rather include regional population among the regressors, see next section.

⁸Because the network includes only those who moved into the region more than 10 years ago, the spatially lagged network is exogenous in the regression, and there is no spatially contemporaneous dependence (i.e., spatial lags or spatial errors, see Anselin, 2006). Besides, there are (to the best knowledge of the authors) no estimators allowing for contemporaneous spatial dependence in the random parameters logit model (see section 3).

⁹Although, for example, Patacchini and Zenou (2008) used networks in neighboring regions as a proxy for weak ties when estimating the effect of networks on the employment probability.

¹⁰Note that if new migrants choose their region randomly (with the same probability for each region) this figure is expected to be as low as 0.6 %.

¹¹See also Bartel (1989), Bauer et al. (2000, 2005, 2007), Gottlieb and Joseph (2006), Jaeger (2007), Christiadi and Cushing (2008) or Grogger and Hanson (2011) for related applications of the conditional logit.

¹²Including the network size of neighboring regions among the regressors is thus also a test for IIA (see also Train, 2009, p. 49).

¹³A probably more common alternative model which relaxes the IIA assumption is the nested logit model. However, while nested logit does not impose IIA between nests, alternatives within a given nest are still assumed to exhibit independence from irrelevant alternatives. The model is thus less flexible than the random parameters logit and therefore not considered.

¹⁴An alternative interpretation of the random parameters logit is based on the error components creating correlations among utilities for different alternatives, which is formally equivalent to this interpretation, see Train (2009), p. 139f.

¹⁵Revelt and Train (1998) and Train (1999) cite Ruud (1996) showing that random parameters logit models have a tendency to be unstable when all coefficients are treated as random. Therefore, at least one coefficient should be fixed.

¹⁶Although heterogeneity of tastes can be expected as regards to individual's preferences for receiving countries, the maximum dimension of the Mata routine to generate the Halton draws in the STATA statistics package (see Drukker and Gates, 2006) is 20, so that no more than 20 unrestricted coefficients in β can be estimated.

¹⁷Although sign restrictions could be imposed by specifying some of the coefficients as being lognormally distributed—for example, the coefficient of income can be expected to be positive for all individuals, although its magnitude may vary between decision makers—we specify the random parameters to be normally distributed to make our model as flexible as possible. Furthermore, lognormal distributions usually have a long right-hand tail, which might be problematic in calculations of the willingness-to-pay or the compensating variation because it often leads to unrealistic mean values (see Hensher and Greene, 2003, for a discussion). The use of the log-normal distribution is also discouraged by Sillano and Ortúzar (2005).

¹⁸Halton sequences are usually defined in terms of a prime number. For the simulation of an integral of dimension ι (where the dimension is equal to the number of random parameters), the first ι prime numbers are conventionally used to create ι sequences (Cappellari and Jenkins, 2006). Because the initial elements of the sequences can be highly correlated across dimensions, Train (2009, p. 227) recommends to discard at least the first κ elements, where κ should be at least as large as the ι th prime number. Because our model uses 16 random parameters, the first 53 elements are dropped.

¹⁹As mentioned in the previous section, source country fixed effects cannot be estimated because they cancel out in logit formula (4).

²⁰The regression results are not reported but available from the authors upon request.

²¹These differences are also substantial within countries: While each region has on average 3.52 (first) neighbors within the same country, only 1.42 of these are within a 100 kilometer radius, 1.30 are within 100–200 kilometers, and 0.80 are more than 200 kilometers away.

²²The 2009 edition of the Human Development Report is used, which reports the HDI based on 2007 figures, see UNDP (2009) for details.

²³The sample is restricted to migrants younger than 55 years of age because older cohorts already contain a large number of retirees. According to 2006 data from the European Labour Force Survey (Eurostat, 2008), the average age at which employed persons started receiving a

retirement pension in the 13 receiving countries considered ranged from 54.5 years in France to 61.7 years in Denmark.

²⁴If migration decisions are made at the household level, Mincer (1978) showed that women are more likely to be “tied movers” while men more often are the “primary movers” of the household. If this is the case or if female household members move later than male household members (for example because family reunion in the host country is not possible immediately after migration), some female migrant’s location choice may depend on the choice made by the husband and not on other factors such as ethnic networks, which may affect the estimated parameters. But as shown by a regression on the restricted sample of male migrants aged 25 to 54 (see table A2 in the appendix) excluding females from the sample has only small quantitative effects on most estimated parameters, so the main conclusions of section 4 remain unchanged.

References

- ÅSLUND, O. (2005): “Now and forever? Initial and subsequent location choices of immigrants,” *Regional Science and Urban Economics*, 35, 141–165.
- ANDERSSON, P. AND E. WADENSJÖ (2009): “The employees of native and immigrant self-employed,” *Research in Labor Economics*, 29, 229–250.
- ANSELIN, L. (2006): “Spatial Econometrics,” in *Palgrave Handbook of Econometrics*, ed. by T. C. Mills and K. Patterson, New York: Palgrave Macmillan, vol. 1, chap. 26, 901–969.
- BARTEL, A. P. (1989): “Where do the new U.S. immigrants live?” *Journal of Labor Economics*, 7, 371–391.
- BAUER, T., G. S. EPSTEIN, AND I. N. GANG (2000): “What are migration networks?” IZA Discussion Paper 200, Institute for the Study of Labour (IZA), Bonn.
- (2005): “Enclaves, language and the location choice of migrants,” *Journal of Population Economics*, 18, 649–662.
- (2007): “The influence of stocks and flows on migrants’ location choices,” *Research in Labor Economics*, 26, 199–229.
- BHAT, C. R. (2001): “Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model,” *Transportation Research B*, 35, 677–693.
- BISIN, A., E. PATACCHINI, T. VERDIER, AND Y. ZENOU (2011): “Ethnic identity and labour market outcomes of immigrants in Europe,” *Economic Policy*, 26, 57–92.

- BORJAS, G. J. (1992): “Ethnic capital and intergenerational mobility,” *The Quarterly Journal of Economics*, 107, 123–150.
- (1995): “Ethnicity, neighborhoods, and human-capital externalities,” *American Economic Review*, 85, 365–390.
- BOYD, J. H. AND R. E. MELLMAN (1980): “The effect of fuel economy standards on the U.S. automotive market: An hedonic demand analysis,” *Transportation Research A*, 14, 367–378.
- CAPPELLARI, L. AND S. P. JENKINS (2006): “Calculation of multivariate normal probabilities by simulation, with applications to maximum simulated likelihood estimation,” *The Stata Journal*, 6, 156–189.
- CARDAK, B. A. AND J. T. McDONALD (2004): “Neighbourhood effects, preference heterogeneity and immigrant educational attainment,” *Applied Economics*, 36, 559–572.
- CARDELL, N. S. AND F. C. DUNBAR (1980): “Measuring the societal impacts of automobile downsizing,” *Transportation Research A*, 14, 423–434.
- CARRINGTON, W. J., E. DETRAGIACHE, AND T. VISHWANATH (1996): “Migration with endogenous moving costs,” *American Economic Review*, 86, 909–930.
- CHISWICK, B. AND P. W. MILLER (2005): “Do enclaves matter in immigrant adjustment?” *City & Community*, 4, 5–35.
- CHRISTIADI AND B. CUSHING (2008): “The joint choice of an individual’s occupation and destination,” *Journal of Regional Science*, 48, 893–919.
- CUTLER, D. M. AND E. L. GLAESER (1997): “Are ghettos good or bad?” *The Quarterly Journal of Economics*, 112, 827–872.
- CUTLER, D. M., E. L. GLAESER, AND J. L. VIGDOR (2008): “Is the Melting Pot still hot? Explaining the resurgence of immigrant segregation,” *The Review of Economics and Statistics*, 90, 478–497.
- DAHLBERG, M. AND M. EKLÖF (2003): “Relaxing the IIA assumption in locational choice models: a comparison between conditional logit, mixed logit, and multinomial probit models,” Working Paper 2003:9, Department of Economics, Uppsala University, Uppsala.
- DAMM, A. P. (2009a): “Determinants of recent immigrants’ location choices: quasi-experimental evidence,” *Journal of Population Economics*, 22, 145–174.

- (2009b): “Ethnic enclaves and immigrant labor market outcomes: quasi-experimental evidence,” *Journal of Labor Economics*, 27, 281–314.
- DAVIES, P. S., M. J. GREENWOOD, AND H. LI (2001): “A conditional logit approach to U.S. state-to-state migration,” *Journal of Regional Science*, 41, 337–360.
- DRUKKER, D. M. AND R. GATES (2006): “Generating Halton sequences using Mata,” *The Stata Journal*, 6, 214–228.
- EDIN, P.-Å., P. FREDRIKSSON, AND O. ÅSLUND (2003): “Ethnic enclaves and the economic success of immigrants – evidence from a natural experiment,” *The Quarterly Journal of Economics*, 118, 329–357.
- EPSTEIN, G. S. (2010): “Informational cascades and the decision to migrate,” in *Migration and Culture*, ed. by G. S. Epstein and I. N. Gang, Bingley: Emerald, chap. 2, 25–44.
- EUROSTAT (2008): *Transition from Work into Retirement*, Luxembourg: Office for Official Publications of the European Communities.
- GEIS, W., S. UEBELMESSER, AND M. WERDING (2008): “How do migrants choose their destination country? An analysis of institutional determinants,” CESifo Working Paper 2506, CESifo, Munich.
- GONZALEZ, A. (1998): “Mexican enclaves and the price of culture,” *Journal of Urban Economics*, 43, 273–291.
- GOTTLIEB, P. D. AND G. JOSEPH (2006): “College-to-work migration of technology graduates and holders of doctorates within the United States,” *Journal of Regional Science*, 46, 627–659.
- GROGGER, J. AND G. HANSON (2011): “Income maximization and the selection and sorting of international migrants,” *Journal of Development Economics*, 95, 42–57.
- GROSS, D. M. AND N. SCHMITT (2003): “The role of cultural clustering in attracting new immigrants,” *Journal of Regional Science*, 43, 295–318.
- HALTON, J. H. (1960): “On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals,” *Numerische Mathematik*, 2, 84–90.
- HEITMUELLER, A. (2006): “Coordination failures in network migration,” *The Manchester School*, 74, 701–710.

- HENSHER, D. A. (2001): “The valuation of commuter travel time savings for car drivers: evaluating alternative model specifications,” *Transportation*, 28, 101–118.
- HENSHER, D. A. AND W. H. GREENE (2003): “The mixed logit model: the state of practice,” *Transportation*, 30, 133–176.
- HOLE, A. R. (2007): “Fitting mixed logit models by using maximum simulated likelihood,” *The Stata Journal*, 7, 388–401.
- JAEGER, D. A. (2007): “Green Cards and the location choices of immigrants in the United States,” *Research in Labor Economics*, 27, 131–183.
- LAZEAR, E. P. (1999): “Culture and language,” *Journal of Political Economy*, 107, 95–126.
- MARSCHAK, J. (1960): “Binary choice constraints on random utility indications,” in *Standord Symposium on Mathematical Methods in the Social Sciences*, ed. by K. Arrow, Stanford: Stanford University Press, 312–322.
- MASSY, D., J. ARANGO, G. HUGO, A. KOUAOUCCI, A. PELLEGRINO, AND J. TAYLOR (1993): “Theories of international migration: a review and appraisal,” *Population and Development Review*, 19, 431–466.
- McFADDEN, D. (1974): “Conditional Logit Analysis of Qualitative Choice Behavior,” in *Frontiers in Econometrics*, ed. by P. Zarembka, New York: Academic Press, chap. 4, 105–142.
- McFADDEN, D. AND K. E. TRAIN (2000): “Mixed MNL models for discrete response,” *Journal of Applied Econometrics*, 15, 447–470.
- MINCER, J. (1978): “Family migration decisions,” *Journal of Political Economy*, 86, 749–773.
- MORETTI, E. (2010): “Local labor markets,” in *Handbook of Labor Economics*, ed. by O. Ashenfelter and D. Card, Amsterdam: North-Holland, vol. 4B, chap. 14, 1237–1313.
- MUNSHI, K. (2003): “Networks in the modern economy: Mexican migrants in the U.S. labor market,” *The Quarterly Journal of Economics*, 118, 549–599.
- ORTEGA, F. AND G. PERI (2009): “The causes and effects of international migrations: evidence from OECD countries 1980–2005,” NBER Working Paper 14833, National Bureau of Economic Research, Cambridge, MA.

- PATACCHINI, E. AND Y. ZENOU (2008): “Ethnic networks and employment outcomes,” IZA Discussion Paper 3331, Institute for the Study of Labour (IZA), Bonn.
- PEDERSEN, P. J., M. PYTLIKOVA, AND N. SMITH (2008): “Selection and network effects—Migration flows into OECD countries 1990–2000,” *European Economic Review*, 52, 1160–1186.
- PORTNOV, B. A. (1999): “The effect of regional inequalities on migration: a comparative analysis of Israel and Japan,” *International Migration*, 37, 587–615.
- REVELT, D. AND K. E. TRAIN (1998): “Mixed logit with repeated choices: households’ choices of appliance efficiency level,” *The Review of Economics and Statistics*, 80, 647–657.
- RUUD, P. A. (1996): “Approximation and simulation of the multinomial probit model: an analysis of covariance matrix estimation,” Working paper, Department of Economics, University of California, Berkeley.
- SILLANO, M. AND J. D. D. ORTÚZAR (2005): “Willingness-to-pay estimation with mixed logit models: some new evidence,” *Environment and Planning A*, 37, 525–550.
- STARK, O. (1994): “Patterns of Labor Migration when Workers Differ in Their Skills,” in *Economic Aspects of International Immigration*, ed. by H. Giersch, Berlin: Springer Verlag.
- TRAIN, K. E. (1999): “Halton sequences for mixed logit,” Working paper, Department of Economics, University of California, Berkeley.
- (2009): *Discrete Choice Methods with Simulation*, New York: Cambridge University Press, second ed.
- UNDP (2009): *Human Development Report 2009. Overcoming Barriers: Human Mobility and Development*, New York: United Nations Development Programme.
- WARNES, T. A. M. (2009): “International retirement migration,” in *International Handbook of Population Aging*, ed. by P. Uhlenberg, Netherlands: Springer, vol. 1 of *International Handbooks of Population*, 341–363.
- ZAVODNY, M. (1999): “Determinants of recent immigrants’ locational choice,” *International Migration Review*, 33, 1014–1030.

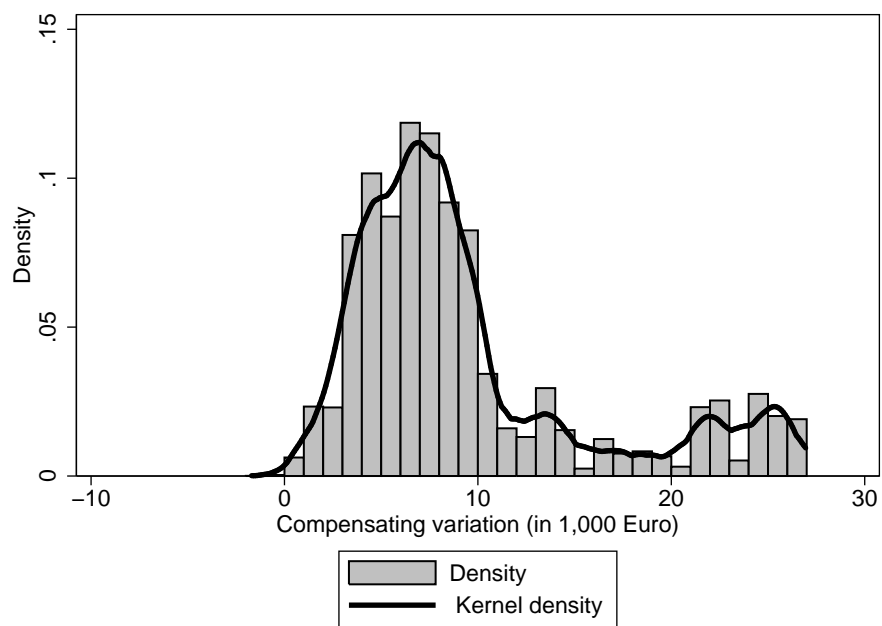


Figure 1: Distribution of compensating variation for a 1 percentage point variation in ethnic network size (in € 1,000). Histogram and kernel density estimate. Calculated from individual level parameters at mean network size. $N = 8,988,710$ observations. Source: European Labour Force Survey 2007.

Variable	Mean	S. D.	Min.	Max.
Relative network $_{js}$	6.650	10.276	0.000	100.000
Relative network $_{js}^{N_1}$	7.449	9.841	0.000	100.000
Relative network $_{js}^{N_2}$	9.571	11.220	0.000	100.000
Relative network $_{js}^{N_C}$	14.442	17.334	0.000	100.000
Relative network $_{js}^{N'_1}$	0.364	1.568	0.000	32.338
Relative network $_{js}^{N'_2}$	1.310	3.570	0.000	51.233
Absolute network $_{js}$ (in 1,000)	14.323	34.175	0.000	265.987
Absolute network $_{js}^{N_1}$ (in 1,000)	16.964	39.340	0.000	463.514
Absolute network $_{js}^{N_2}$ (in 1,000)	22.496	49.831	0.000	652.008
Absolute network $_{js}^{N_C}$ (in 1,000)	35.114	79.183	0.000	857.423
Absolute network $_{js}^{N'_1}$ (in 1,000)	2.564	13.955	0.000	365.762
Absolute network $_{js}^{N'_2}$ (in 1,000)	8.085	31.300	0.000	411.412

Table 1: Summary statistics, network size variables. $N = 8,988,710$ observations. Source: European Labour Force Survey 2007. N_1 : set of neighboring regions within the host country. N_2 : set of second neighbor regions within the host country. N_C : set of all other regions in the host country. N'_1 : set of neighboring regions in other countries. N'_2 : set of second neighbor regions in other countries.

Variable	Mean	S. D.	Min.	Max.
Population (in 100,000) [†]	1.544	1.449	0.107	9.027
Region size (in 1,000 km ²) [†]	17.345	23.686	0.161	165.296
Unemployment rate (in %) [†]	7.290	3.743	2.286	20.186
Avg. income p. a. (in € 1,000) [†]	27.263	10.299	10.567	95.979
Capital (= 1) [†]	0.082	0.275	0.000	1.000
Distance (in 1,000 km)	4.697	3.641	0.055	18.981
Common border (= 1)	0.045	0.207	0.000	1.000
Common official language (= 1)	0.375	0.484	0.000	1.000
Colony after 1945 (= 1)	0.140	0.347	0.000	1.000

Table 2: Summary statistics, region specific variables ([†] $N = 158$ observations) and country-pair specific variables ($N = 8,988,710$ observations). Source: Eurostat, CEPII.

Variable	RPL			CL	
	Mean(β)	S. D.(β)	% $\beta > 0$	$e^{\text{Mean}(\beta)}$	β e^β
Relative network $_{js}$	0.376*** (0.001)	0.014*** (0.000)	100.000	1.456*** (0.001)	0.127*** (0.000)
(Relative network $_{js}$) ²	-0.017*** (0.000)	0.013*** (0.000)	10.229	0.983*** (0.000)	-0.001*** (0.000)
Relative network $_{js}^{N_1}$	0.051*** (0.000)	0.001*** (0.000)	100.000	1.053*** (0.000)	0.049*** (0.000)
Relative network $_{js}^{N_2}$	0.036*** (0.000)	0.004*** (0.000)	100.000	1.037*** (0.000)	0.031*** (0.000)
Relative network $_{js}^{N_C}$	0.025*** (0.000)	0.034*** (0.000)	76.856	1.025*** (0.000)	0.026*** (0.000)
Relative network $_{js}^{N'_1}$	0.034*** (0.000)	0.002*** (0.001)	100.000	1.035*** (0.000)	0.043*** (0.000)
Relative network $_{js}^{N'_2}$	0.012*** (0.000)	0.005*** (0.000)	98.720	1.012*** (0.000)	0.014*** (0.000)
Population (in 100,000)	0.290*** (0.000)	0.001*** (0.000)	100.000	1.336*** (0.000)	0.324*** (0.000)
Region size (in 1,000 km ²)	-0.005*** (0.000)			0.995*** (0.000)	-0.004*** (0.000)
Unemployment rate (in %)	-0.022*** (0.000)	0.005*** (0.001)	0.001	0.978*** (0.000)	-0.019*** (0.000)
Avg. income p. a. (in € 1,000)	0.016*** (0.000)	0.001** (0.000)	100.000	1.016*** (0.000)	0.017*** (0.000)
Distance (in 1,000 km)	-0.203*** (0.001)	0.073*** (0.005)	0.280	0.816*** (0.001)	-0.412*** (0.001)
Distance (in 1,000 km) ²	0.012*** (0.000)	0.005*** (0.000)	99.523	1.012*** (0.000)	0.023*** (0.000)
Capital (= 1)	-1.731*** (0.008)	3.297*** (0.010)	29.972	0.177*** (0.001)	-0.004*** (0.001)
Common border (= 1)	-0.264*** (0.007)	2.716*** (0.011)	46.135	0.768*** (0.005)	0.524*** (0.003)
Common official language (= 1)	1.088*** (0.002)	1.587*** (0.005)	75.355	2.968*** (0.007)	0.863*** (0.002)
Colony after 1945 (= 1)	-1.572*** (0.004)	1.280*** (0.009)	10.979	0.208*** (0.001)	-1.036*** (0.003)
Observations			8,988,710		8,988,710

Table 3: Random parameters logit (RPL) and conditional logit (CL) regressions of location choice using relative network size. Germany and Ireland not included. Receiving country fixed effects reported in table A1. Standard errors in parentheses. *** significant at 1 %, ** significant at 5 %, * significant at 10 %. RPL log likelihood simulated using 500 Halton draws. Source: European Labour Force Survey 2007, Eurostat, CEPII. N_1 : set of neighboring regions within the host country. N_2 : set of second neighbor regions within the host country. N_C : set of all other regions in the host country. N'_1 : set of neighboring regions in other countries. N'_2 : set of second neighbor regions in other countries.

Variable	Mean	S. D.	Min.	Max.
$CV_k(\text{Relative network}_{js} = 1)$	21.612	0.982	19.734	24.290
$CV_k(\text{Relative network}_{js} = 5)$	13.010	4.809	4.564	26.170
$CV_k(\text{Relative network}_{js} = 10)$	2.257	9.594	-14.415	28.519
$CV_k(\text{Relative network}_{js} = 15)$	-8.495	14.379	-33.394	30.868
$CV_k(\text{Relative network}_{js} = 20)$	-19.248	19.164	-52.373	33.218
$CV_k(\text{Relative network}_{js} = 25)$	-30.000	23.949	-71.352	35.567
$CV_k\left(\text{Relative network}_{js}^{N_1}\right)$	3.245	0.005	3.184	3.285
$CV_k\left(\text{Relative network}_{js}^{N_2}\right)$	2.286	0.026	2.128	2.463
$CV_k\left(\text{Relative network}_{js}^{N_C}\right)$	1.558	0.677	-0.535	5.387
$CV_k\left(\text{Relative network}_{js}^{N'_1}\right)$	2.182	0.009	2.124	2.244
$CV_k\left(\text{Relative network}_{js}^{N'_2}\right)$	0.766	0.024	0.613	1.011

Table 4: Compensating variation (willingness to pay) for a 1 percentage point change in network size (in € 1,000). $N = 8,988,710$ observations. Source: European Labour Force Survey 2007.

Variable	RPL			CL	
	Mean(β)	S. D. (β)	% $\beta > 0$	$e^{\text{Mean}(\beta)}$	β e^β
Absolute Network k_{js} (in 1,000)	0.287*** (0.000)	0.035*** (0.000)	100.000	1.332*** (0.000)	0.045*** (0.000)
(Absolute Network k_{js}) ² (in 1,000)	-0.006*** (0.000)	0.005*** (0.000)	11.132	0.994*** (0.000)	-0.000*** (0.000)
Absolute Network $k_{js}^{N_1}$ (in 1,000)	0.002*** (0.000)	0.000*** (0.000)	100.000	1.002*** (0.000)	0.003*** (0.000)
Absolute Network $k_{js}^{N_2}$ (in 1,000)	0.001*** (0.000)	0.000*** (0.000)	100.000	1.001*** (0.000)	0.001*** (0.000)
Absolute Network $k_{js}^{N_C}$ (in 1,000)	-0.001*** (0.000)	0.002*** (0.000)	36.764	0.999*** (0.000)	0.001*** (0.000)
Absolute Network $k_{js}^{N'_1}$ (in 1,000)	0.005*** (0.000)	0.002*** (0.000)	99.307	1.005*** (0.000)	0.002*** (0.000)
Absolute Network $k_{js}^{N'_2}$ (in 1,000)	0.002*** (0.000)	0.002*** (0.000)	80.126	1.002*** (0.000)	0.001*** (0.000)
Population (in 100,000)	0.250*** (0.000)	0.003*** (0.001)	100.000	1.284*** (0.000)	0.357*** (0.000)
Region size (in 1,000 km ²)	-0.001*** (0.000)			0.999*** (0.000)	-0.004*** (0.000)
Unemployment rate (in %)	-0.037*** (0.000)	0.035*** (0.001)	14.530	0.964*** (0.000)	-0.037*** (0.000)
Avg. income p. a. (in € 1,000)	0.014*** (0.000)	0.014*** (0.000)	83.971	1.014*** (0.000)	0.021*** (0.000)
Distance (in 1,000 km)	-0.474*** (0.001)	0.007*** (0.003)	0.000	0.622*** (0.001)	-0.610*** (0.001)
Distance (in 1,000 km) ²	0.021*** (0.000)	0.000*** (0.000)	100.000	1.021*** (0.000)	0.022*** (0.000)
Capital (= 1)	-5.436*** (0.025)	8.016*** (0.033)	24.886	0.004*** (0.000)	0.105*** (0.001)
Common border (= 1)	0.464*** (0.005)	1.709*** (0.008)	60.697	1.590*** (0.008)	0.958*** (0.003)
Common official language (= 1)	1.689*** (0.002)	0.080*** (0.004)	100.000	5.416*** (0.010)	1.784*** (0.001)
Colony after 1945 (= 1)	-0.243*** (0.004)	1.059*** (0.007)	40.917	0.784*** (0.003)	0.066*** (0.003)
Observations	8,988,710			8,988,710	

Table 5: Random parameters logit (RPL) and conditional logit (CL) regressions of location choice using absolute network size. Germany and Ireland not included. Receiving country fixed effects reported in table A1. Standard errors in parentheses. *** significant at 1 %, ** significant at 5 %, * significant at 10 %. RPL log likelihood simulated using 500 Halton draws. Source: European Labour Force Survey 2007, Eurostat, CEPII. N_1 : set of neighboring regions within the host country. N_2 : set of second neighbor regions within the host country. N_C : set of all other regions in the host country. N'_1 : set of neighboring regions in other countries. N'_2 : set of second neighbor regions in other countries.

Variable	Mean	S. D.	Min.	Max.
Relative network $_j^{\rho \leq 100}$	4.008	9.965	0.000	97.181
Relative network $_j^{100 < \rho \leq 200}$	5.106	9.777	0.000	100.000
Relative network $_j^{\rho > 200}$	21.953	21.113	0.000	100.000
Relative network $_j^{\rho' \leq 100}$	0.061	0.857	0.000	60.556
Relative network $_j^{100 < \rho' \leq 200}$	0.469	2.443	0.000	89.838

Table 6: Summary statistics, networks in neighboring regions defined by distance from region of residence. $N = 8,988,710$ observations. Source: European Labour Force Survey 2007.

Variable	RPL			CL	
	Mean(β)	S. D.(β)	% $\beta > 0$	$e^{\text{Mean}(\beta)}$	β e^β
Relative network _{js}	0.450*** (0.001)	0.001*** (0.000)	100.000	1.568*** (0.001)	0.135*** (0.000)
(Relative network _{js}) ²	-0.024*** (0.000)	0.019*** (0.000)	10.017	0.976*** (0.000)	-0.001*** (0.000)
Relative network _j ^{$\rho \leq 100$}	0.041*** (0.000)	0.005*** (0.000)	100.000	1.042*** (0.000)	0.042*** (0.000)
Relative network _j ^{$100 < \rho \leq 200$}	0.036*** (0.000)	0.003*** (0.000)	100.000	1.037*** (0.000)	0.035*** (0.000)
Relative network _j ^{$\rho > 200$}	0.033*** (0.000)	0.001*** (0.000)	100.000	1.033*** (0.000)	0.030*** (0.000)
Relative network _j ^{$\rho' \leq 100$}	-0.099*** (0.002)	0.098*** (0.001)	15.750	0.906*** (0.001)	0.005*** (0.001)
Relative network _j ^{$100 < \rho' \leq 200$}	0.023*** (0.000)	0.001 (0.000)	100.000	1.024*** (0.000)	0.028*** (0.000)
Population (in 100,000)	0.278*** (0.000)	0.006*** (0.000)	100.000	1.321*** (0.000)	0.317*** (0.000)
Region size (in 1,000 km ²)	-0.002*** (0.000)			0.998*** (0.000)	-0.002*** (0.000)
Unemployment rate (in %)	-0.039*** (0.000)	0.010*** (0.001)	0.002	0.961*** (0.000)	-0.034*** (0.000)
Avg. income p.a. (in € 1,000)	0.015*** (0.000)	0.010*** (0.000)	91.999	1.015*** (0.000)	0.019*** (0.000)
Distance (in 1,000 km)	-0.212*** (0.001)	0.008* (0.004)	0.000	0.809*** (0.001)	-0.463*** (0.001)
Distance (in 1,000 km) ²	0.011*** (0.000)	0.003*** (0.000)	99.936	1.011*** (0.000)	0.023*** (0.000)
Capital (= 1)	-1.904*** (0.007)	3.384*** (0.009)	28.679	0.149*** (0.001)	-0.095*** (0.001)
Common border (= 1)	-0.138*** (0.007)	2.493*** (0.010)	47.795	0.871*** (0.006)	0.543*** (0.003)
Common official language (= 1)	1.081*** (0.002)	1.642*** (0.005)	74.480	2.946*** (0.007)	0.820*** (0.002)
Colony after 1945 (= 1)	-1.698*** (0.004)	1.587*** (0.007)	14.227	0.183*** (0.001)	-1.067*** (0.003)
Observations			8,988,710		8,988,710

Table 7: Random parameters logit (RPL) and conditional logit (CL) regressions of location choice using relative network size. Germany and Ireland not included. Receiving country fixed effects reported in table A1. Standard errors in parentheses. *** significant at 1 %, ** significant at 5 %, * significant at 10 %. RPL log likelihood simulated using 500 Halton draws. Source: European Labour Force Survey 2007, Eurostat, CEPII.

Variable	RPL			CL	
	Mean(β)	S. D.(β)	% $\beta > 0$	$e^{\text{Mean}(\beta)}$	β e^β
Relative network _{<i>js</i>}	0.427*** (0.001)	0.042*** (0.001)	100.000	1.533*** (0.001)	0.132*** (0.000)
(Relative network _{<i>js</i>}) ²	-0.023*** (0.000)	0.018*** (0.000)	10.395	0.978*** (0.000)	-0.002*** (0.000)
Relative network _{<i>js</i>} ^{<i>N</i>₁}	0.052*** (0.000)	0.001*** (0.000)	100.000	1.054*** (0.000)	0.050*** (0.000)
Relative network _{<i>js</i>} ^{<i>N</i>₂}	0.036*** (0.000)	0.012*** (0.000)	99.881	1.037*** (0.000)	0.030*** (0.000)
Relative network _{<i>js</i>} ^{<i>N</i>_C}	0.025*** (0.000)	0.036*** (0.000)	76.131	1.026*** (0.000)	0.027*** (0.000)
Relative network _{<i>js</i>} ^{<i>N</i>₁'}	0.029*** (0.000)	0.019*** (0.001)	93.831	1.029*** (0.000)	0.041*** (0.000)
Relative network _{<i>js</i>} ^{<i>N</i>₂'}	0.011*** (0.000)	0.005*** (0.000)	98.060	1.011*** (0.000)	0.012*** (0.000)
Population (in 100,000)	0.284*** (0.000)	0.003*** (0.000)	100.000	1.328*** (0.000)	0.317*** (0.000)
Region size (in 1,000 km ²)	-0.004*** (0.000)	0.006*** (0.001)	0.008	0.996*** (0.000)	-0.004*** (0.000)
Unemployment rate (in %)	-0.023*** (0.000)	0.006*** (0.001)	0.008	0.977*** (0.000)	-0.022*** (0.000)
Avg. income p.a. (in € 1,000)	0.016*** (0.000)	0.005*** (0.000)	99.894	1.016*** (0.000)	0.018*** (0.000)
Distance (in 1,000 km)	-0.168*** (0.001)	0.261*** (0.004)	25.948	0.845*** (0.001)	-0.401*** (0.001)
Distance (in 1,000 km) ²	0.009*** (0.000)	0.004*** (0.000)	99.258	1.009*** (0.000)	0.021*** (0.000)
Capital (= 1)	-2.024*** (0.008)	3.636*** (0.011)	28.892	0.132*** (0.001)	0.022*** (0.001)
Common border (= 1)	-0.169*** (0.007)	2.561*** (0.010)	47.370	0.845*** (0.006)	0.538*** (0.003)
Common official language (= 1)	1.065*** (0.003)	1.554*** (0.006)	75.353	2.901*** (0.007)	0.834*** (0.002)
Colony after 1945 (= 1)	-1.792*** (0.005)	1.250*** (0.009)	7.584	0.167*** (0.001)	-1.115*** (0.003)
Observations	8,506,229			8,506,229	

Table 8: Random parameters logit (RPL) and conditional logit (CL) regressions of location choice using relative network size. Sample restricted to migrants from countries with at least “medium” human development (HDI ≥ 0.5 , see UNDP, 2009). Germany and Ireland not included. Receiving country fixed effects reported in table A1. Standard errors in parentheses. *** significant at 1 %, ** significant at 5 %, * significant at 10 %. RPL log likelihood simulated using 500 Halton draws. Source: European Labour Force Survey 2007, Eurostat, CEPIL. N_1 : set of neighboring regions within the host country. N_2 : set of second neighbor regions within the host country. N_C : set of all other regions in the host country. N_1' : set of neighboring regions in other countries. N_2' : set of second neighbor regions in other countries.

Variable	RPL			CL	
	Mean(β)	S. D.(β)	% $\beta > 0$	$e^{\text{Mean}(\beta)}$	β e^β
Relative network _{js}	0.410*** (0.001)	0.026*** (0.001)	100.000	1.507*** (0.001)	0.124*** (0.000)
(Relative network _{js}) ²	-0.021*** (0.000)	0.016*** (0.000)	9.415	0.979*** (0.000)	-0.001*** (0.000)
Relative network _{js} ^{N₁}	0.050*** (0.000)	0.002*** (0.000)	100.000	1.051*** (0.000)	0.050*** (0.000)
Relative network _{js} ^{N₂}	0.033*** (0.000)	0.016*** (0.000)	98.249	1.034*** (0.000)	0.031*** (0.000)
Relative network _{js} ^{N_C}	0.023*** (0.000)	0.034*** (0.000)	75.038	1.023*** (0.000)	0.026*** (0.000)
Relative network _{js} ^{N'₁}	0.022*** (0.000)	0.023*** (0.001)	82.607	1.022*** (0.000)	0.038*** (0.000)
Relative network _{js} ^{N'₂}	0.016*** (0.000)	0.002*** (0.000)	100.000	1.017*** (0.000)	0.017*** (0.000)
Population (in 100,000)	0.294*** (0.000)	0.000 (0.001)	100.000	1.341*** (0.000)	0.329*** (0.000)
Region size (in 1,000 km ²)	-0.005*** (0.000)			0.995*** (0.000)	-0.005*** (0.000)
Unemployment rate (in %)	-0.021*** (0.000)	0.015*** (0.001)	8.821	0.979*** (0.000)	-0.017*** (0.000)
Avg. income p.a. (in € 1,000)	0.015*** (0.000)	0.006*** (0.000)	99.265	1.015*** (0.000)	0.018*** (0.000)
Distance (in 1,000 km)	-0.180*** (0.002)	0.058*** (0.007)	0.101	0.836*** (0.001)	-0.393*** (0.001)
Distance (in 1,000 km) ²	0.010*** (0.000)	0.002*** (0.000)	99.999	1.010*** (0.000)	0.022*** (0.000)
Capital (= 1)	-1.668*** (0.010)	3.218*** (0.013)	30.214	0.189*** (0.002)	0.005*** (0.002)
Common border (= 1)	-0.170*** (0.008)	2.421*** (0.011)	47.207	0.844*** (0.006)	0.507*** (0.003)
Common official language (= 1)	1.087*** (0.003)	1.313*** (0.006)	79.604	2.965*** (0.008)	0.870*** (0.002)
Colony after 1945 (= 1)	-1.589*** (0.005)	1.484*** (0.009)	14.202	0.204*** (0.001)	-1.037*** (0.003)
Observations			6,732,410		6,732,410

Table 9: Random parameters logit (RPL) and conditional logit (CL) regressions of location choice using relative network size. Sample restricted to migrants age 25 to 54. Germany and Ireland not included. Receiving country fixed effects reported in table A1. Standard errors in parentheses. *** significant at 1 %, ** significant at 5 %, * significant at 10 %. RPL log likelihood simulated using 500 Halton draws. Source: European Labour Force Survey 2007, Eurostat, CEPIL. N_1 : set of neighboring regions within the host country. N_2 : set of second neighbor regions within the host country. N_C : set of all other regions in the host country. N'_1 : set of neighboring regions in other countries. N'_2 : set of second neighbor regions in other countries.

	Table 3	Table 5	Table 7	Table 8	Table 9	Table A2
Belgium (=1)	-0.335*** (0.004)	-0.294*** (0.003)	-0.288*** (0.004)	-0.325*** (0.004)	-0.387*** (0.004)	-0.439*** (0.006)
Denmark (=1)	0.532*** (0.005)	0.404*** (0.005)	0.534*** (0.005)	0.285*** (0.005)	0.512*** (0.006)	0.410*** (0.009)
Spain (=1)	1.449*** (0.003)	1.744*** (0.003)	1.556*** (0.003)	1.530*** (0.003)	1.500*** (0.003)	1.602*** (0.005)
Finland (=1)	-1.529*** (0.011)	-1.370*** (0.009)	-1.405*** (0.011)	-1.471*** (0.011)	-1.399*** (0.013)	-1.249*** (0.021)
France (=1)	-0.964*** (0.003)	-0.824*** (0.003)	-0.838*** (0.003)	-0.838*** (0.003)	-0.958*** (0.004)	-0.991*** (0.005)
Greece (=1)	-0.609*** (0.004)	-0.334*** (0.004)	-0.512*** (0.005)	-0.555*** (0.005)	-0.562*** (0.005)	-0.325*** (0.007)
Italy (=1)	0.188*** (0.003)	0.657*** (0.003)	0.223*** (0.003)	0.258*** (0.003)	0.237*** (0.003)	0.128*** (0.005)
Luxembourg (=1)	-0.889*** (0.009)	-1.122*** (0.009)	-0.883*** (0.009)	-0.842*** (0.010)	-0.721*** (0.010)	-0.845*** (0.017)
Netherlands (=1)	-0.405*** (0.003)	-0.082*** (0.003)	-0.395*** (0.003)	-0.317*** (0.003)	-0.295*** (0.004)	-0.389*** (0.006)
Portugal (=1)	0.361*** (0.004)	0.437*** (0.004)	0.469*** (0.004)	0.433*** (0.004)	0.432*** (0.005)	0.659*** (0.007)
Sweden (=1)	0.129*** (0.003)	0.105*** (0.003)	0.141*** (0.004)	-0.006 (0.004)	0.190*** (0.004)	0.230*** (0.006)
United Kingdom (=1)	0.481*** (0.003)	0.763*** (0.003)	0.431*** (0.003)	0.498*** (0.003)	0.465*** (0.003)	0.638*** (0.005)

Table A1: Receiving country fixed effects, random parameters logit regressions. Base category: Austria. Germany and Ireland not included. Standard errors in parentheses. *** significant at 1 %, ** significant at 5 %, * significant at 10 %. RPL log likelihood simulated using 500 Halton draws. Source: European Labour Force Survey 2007, Eurostat, CEPII.

Variable	RPL			CL	
	Mean(β)	S. D.(β)	% $\beta > 0$	$e^{\text{Mean}(\beta)}$	β e^β
Relative network _{js}	0.341*** (0.001)	0.024*** (0.001)	100.000	1.406*** (0.001)	0.113*** (0.000)
(Relative network _{js}) ²	-0.015*** (0.000)	0.011*** (0.000)	9.261	0.985*** (0.000)	-0.001*** (0.000)
Relative network _{js} ^{N₁}	0.057*** (0.000)	0.006*** (0.000)	100.000	1.058*** (0.000)	0.053*** (0.000)
Relative network _{js} ^{N₂}	0.041*** (0.000)	0.007*** (0.000)	100.000	1.042*** (0.000)	0.035*** (0.000)
Relative network _{js} ^{N_C}	0.030*** (0.000)	0.029*** (0.000)	84.604	1.030*** (0.000)	0.028*** (0.000)
Relative network _{js} ^{N'₁}	-0.001** (0.001)	0.066*** (0.001)	49.100	0.999** (0.001)	0.036*** (0.000)
Relative network _{js} ^{N'₂}	0.022*** (0.000)	0.008*** (0.000)	99.625	1.022*** (0.000)	0.022*** (0.000)
Population (in 100,000)	0.324*** (0.000)	0.006*** (0.001)	100.000	1.383*** (0.001)	0.335*** (0.000)
Region size (in 1,000 km ²)	-0.008*** (0.000)			0.992*** (0.000)	(0.000) (0.001)
Unemployment rate (in %)	-0.014*** (0.000)	0.001** (0.001)	0.000	0.986*** (0.000)	-0.014*** (0.000)
Avg. income p.a. (in € 1,000)	0.013*** (0.000)	0.015*** (0.000)	81.480	1.013*** (0.000)	0.017*** (0.000)
Distance (in 1,000 km)	-0.205*** (0.002)	0.110*** (0.007)	3.134	0.815*** (0.002)	-0.385*** (0.001)
Distance (in 1,000 km) ²	0.010*** (0.000)	0.005*** (0.001)	97.856	1.010*** (0.000)	0.019*** (0.000)
Capital (= 1)	-1.755*** (0.012)	3.249*** (0.016)	29.455	0.173*** (0.002)	0.038*** (0.002)
Common border (= 1)	-0.658*** (0.013)	3.025*** (0.018)	41.389	0.518*** (0.007)	0.408*** (0.005)
Common official language (= 1)	1.097*** (0.004)	1.478*** (0.008)	77.093	2.995*** (0.013)	0.860*** (0.003)
Colony after 1945 (= 1)	-1.918*** (0.007)	1.271*** (0.012)	6.586	0.147*** (0.001)	-1.277*** (0.005)
Observations			3,177,240		3,177,240

Table A2: Random parameters logit (RPL) and conditional logit (CL) regressions of location choice using relative network size. Sample restricted to male migrants age 25 to 54. Germany and Ireland not included. Receiving country fixed effects reported in table A1. Standard errors in parentheses. *** significant at 1 %, ** significant at 5 %, * significant at 10 %. RPL log likelihood simulated using 500 Halton draws. Source: European Labour Force Survey 2007, Eurostat, CEPII. N_1 : set of neighboring regions within the host country. N_2 : set of second neighbor regions within the host country. N_C : set of all other regions in the host country. N'_1 : set of neighboring regions in other countries. N'_2 : set of second neighbor regions in other countries.

Working Papers in Economics and Finance

University of Salzburg

- 2012-07 **Klaus Nowotny** and **Dieter Pennerstorfer**. Ethnic Networks and the Location Choice of Migrants in Europe.
- 2012-06 **Benjamin Furlan**, **Martin Gächter**, **Bob Krebs** and **Harald Oberhofer**. Democratization and Real Exchange Rates.
- 2012-05 **Peter Huber**, **Harald Oberhofer** and **Michael Pfaffermayr**. Job Creation and the Intra-distribution Dynamics of the Firm Size Distribution.
- 2012-04 **Jörg Paetzold**. The Convergence of Welfare State Indicators in Europe: Evidence from Panel Data.
- 2012-03 **Sebastian Rathner**. The Performance of Socially Responsible Investment Funds: A Meta-Analysis.
- 2012-02 **Jesus Crespo Cuaresma** and **Matthias Stöckl**. The Effect of Marketing Spending on Sales in the Premium Car Segment: New Evidence from Germany.
- 2012-01 **Harald Oberhofer**, **Matthias Stöckl** and **Hannes Winner**. The Wage Premium of Globalization: Evidence from European Mergers and Acquisitions.
-
- 2011-06 **Peter Huber**. The self-selection of Commuters.
- 2011-05 **Martin Gächter**, **Peter Schwazer**, **Engelbert Theurl** and **Hannes Winner**. Physician Density in a Two-Tiered Health Care System. Forthcoming in *Health Policy*.
- 2011-04 **Jesús Crespo Cuaresma** and **Max Roser**. Borders Redrawn: Measuring the Statistical Creation of International Trade. Forthcoming in *The World Economy*.
- 2011-03 **Harald Oberhofer** and **Michael Pfaffermayr**. FDI versus Exports: Multiple Host Countries and Empirical Evidence. Published in *The World Economy*.
- 2011-02 **Andrea M. Leiter**, **Magdalena Thöni** and **Hannes Winner**. Duo Cum Faciunt Idem, Non Est Idem: Evidence from Austrian Pain and Suffering Verdicts. Forthcoming as Pricing damages for pain and suffering in courts: The impact of the valuation method in *Journal of Empirical Legal Studies*.
- 2011-01 **Harald Oberhofer** and **Michael Pfaffermayr**. Testing the One-Part Fractional Response Model against an Alternative Two-Part Model.
-
- 2010-16 **Harald Oberhofer**, **Tassilo Philippovich** and **Hannes Winner**. Firm Survival in Professional Football: Evidence from the German Football League.

- 2010-15 **Engelbert Theurl** and **Hannes Winner**. The Male-Female Gap in Physician Earnings: Evidence from a Public Health Insurance System. Published in *Health Economics*.
- 2010-14 **Martin Feldkircher**. Forecast Combination and Bayesian Model Averaging - A Prior Sensitivity Analysis. Published in *Journal of Forecasting*.
- 2010-13 **Jesús Crespo Cuaresma** and **Octavio Fernández Amador**. Business Cycle Convergence in EMU: A Second Look at the Second Moment.
- 2010-12 **Martin Feldkircher** and **Stefan Zeugner**. The Impact of Data Revisions on the Robustness of Growth Determinants - A Note on 'Determinants of Economic Growth: Will Data Tell?'. Published in *Journal of Applied Econometrics*.
- 2010-11 **Andrea M. Leiter**, **Magdalena Thöni** and **Hannes Winner**. Evaluating Human Life Using Court Decisions on Damages for Pain and Suffering. Forthcoming in *International Review of Law and Economics*.
- 2010-10 **Harald Oberhofer**. Employment Effects of Acquisitions: Evidence from Acquired European Firms. Forthcoming in *Review of Industrial Organization*.
- 2010-09 **Christian Reiner**. Regionale Arbeitsmärkte in der „Großen Rezession“: Dynamik regionaler Arbeitslosenquoten in Deutschland, Frankreich und Großbritannien im Krisenjahr 2009. Published in *Zeitschrift für Wirtschaftsgeographie*.
- 2010-08 **Leonardo Baccini** and **Andreas Dür**. The New Regionalism and Policy Interdependence. Published in *British Journal of Political Science*.
- 2010-07 **Harald Oberhofer** and **Michael Pfaffermayr**. Firm Growth in Multinational Corporate Groups. Forthcoming in *Empirical Economics*.
- 2010-06 **Sven P. Jost**, **Michael Pfaffermayr** and **Hannes Winner**. Transfer Pricing as a Tax Compliance Risk.
- 2010-05 **Christian Reiner**. Selling the Ivory Tower and Regional Development: Technology Transfer Offices as Mediators of University-Industry Linkages. Published as University policy and regional development: Technology transfer offices as facilitators and generators of university-industry linkages in *Berichte zur Deutschen Landeskunde*.
- 2010-04 **Matthias Stöckl**. Fremdkapitalquoten in Europa: Ein Ländervergleich. Published in *Wirtschaftspolitische Blätter*.
- 2010-03 **Jesús Crespo Cuaresma**, **Harald Oberhofer** and **Paul A. Raschky**. Oil and the Duration of Dictatorships. Published in *Public Choice*.
- 2010-02 **Matthias Stöckl** and **Hannes Winner**. Körperschaftsbesteuerung und Unternehmensverschuldung: Empirische Evidenz von europäischen Firmendaten.
- 2010-01 **Andrea M. Leiter**, **Andrea M. Parolini** and **Hannes Winner**. Environmental Regulation and Investment: Evidence from European Country-Industry Data. Published in *Ecological Economics*
-

- 2009-06 **Sven P. Jost**. Transfer Pricing Risk Awareness of Multinational Corporations: Evidence from a Global Survey.
- 2009-05 **Hannes Winner**. Der Kampf gegen internationale Steuerhinterziehung: Die OECD Initiativen gegen “Steueroasen”. Published in *Steuer und Wirtschaft*.
- 2009-04 **Michael Pfaffermayr**, **Matthias Stöckl** and **Hannes Winner**. Capital Structure, Corporate Taxation and Firm Age.
- 2009-03 **Simon Loretz** and **Padraig J. Moore**. Corporate Tax Competition Between Firms.
- 2009-02 **Ronald W. McQuaid** and **Walter Scherrer**. Changing Reasons for Public Private Partnerships. Published in *Public Money and Management*.
- 2009-01 **Harald Oberhofer**, **Tassilo Philippovich** and **Hannes Winner**. Distance Matters in Away Games: Evidence from the German Football League. Published in *Journal of Economic Psychology*.