CERGE-EI

Center for Economic Research and Graduate Education

Charles University Prague

Essays on the Economics of Labor Migration

Dmytro Vikhrov

Dissertation

Prague, June 2015

Dissertation Committee:

Byeongju Jeong (CERGE-EI, chair)

Randall K. Filer (Hunter College, CUNY)

Jan Kmenta (University of Michigan)

Nikolas Mittag (CERGE-EI)

Evangelia Vourvachaki (Central Bank of Greece)

Referees:

Sergio Destefanis (Università Degli Studi di Salerno)

Silke Übelmesser (Friedrich-Schiller-Universität Jena)

Acknowledgements

I would like to thank my family for supporting the decision to enroll in a Ph.D. program. I am grateful to CERGE-EI for the financial support provided. Without these two crucial components I would not have been able to start my Ph.D. studies in the first place.

I am indebted to my supervisor Byeongju Jeong for thorough guidance in preparing my dissertation proposal and writing my first ever working paper. I am also grateful to Jan Kmenta, Randy Filer, Nikolas Mittag and Sherzod Tashpulatov for their comments and support. I thank Andrea Downing and Paul Whitaker for English editing.

I am further thankful to Olexandr Nikolaychuk and Simon Commander with whom we did a challenging project on emigration from Ukraine and which later turned into Chapter three of this thesis. I am also grateful to Robert Stojanov with whom I collaborated during my employment in the Czech Academy of Sciences. I am glad that one of our projects successfully turned into a published paper.

Finally, I would like to thank many various people who I met at numerous conferences, seminars and summer schools. Many of their presentations, comments, questions and even simple chats proved to be so useful in shaping my vision.

Introduction

This dissertation consists of four chapters, which I wrote during my Ph.D. studies at CERGE-EI. Being a migrant, I felt that my personal migration experience had something unique to share and this is how the topic of my thesis emerged.

All four chapters analyze labor migration from different perspectives. In the first two chapters I research immigration policy. In a dynamic world where new technologies rapidly reduce mobility costs, immigration policy becomes an important tool in controlling immigration. In the remaining two chapters I focus on the issues of self-selection into emigration using the example of Ukraine and within-country mobility using the example of the Czech Republic. These patterns are important because they determine the direction and magnitude of welfare changes initiated by the mobility of labor.

In Chapter one, I develop a theoretical model, which explains why and when a country imposes entry restrictions on the number and skill type of foreign workers. By imposing an immigration quota, a destination country indirectly affects the welfare of the origin country. Under some conditions, the quota positively affects the sending country welfare because it reduces the extent of the downward effect of new migrants on the wages at the destination. Further, I describe how the quota changes when two countries form a political union.

In Chapter two, I construct an immigration policy index which is a proxy for the laxity of immigration policy. This index has several advantages over existing measures. It is defined and comparable for all countries in the world, varies across destination-origin country pairs and over time. When I use this index in estimation, it accounts for a significant share of migrants in stock data. It also explains gender and education composition of migrant labor.

In Chapter three, I research the selection patterns of migrants from my home country, Ukraine. We conducted a large-scale survey, in which we collected information on migrants' observable characteristics and their labor market outcomes before and after emigration. Using this dataset, we find that Ukrainian migrants are positively selected in terms of age, education, and pre-migration income. This, however, is not reflected in their labor market outcomes because many of them are employed in occupations below their reported education levels. This may be understood in terms of strict immigration policies, high search costs, poor transferability of human capital obtained in Ukraine or individual unobservable skills.

In Chapter four, I research selection into internal migration in the East of the Czech Republic. This part of the country is constantly subject to relatively high flood risks from nearby water sources. To cover flood related losses and reduce household vulnerability, many people start commuting for work to nearby larger cities, which offer better employment opportunities and higher pay. Interestingly, the surveyed area is characterized by a high level of permanent out-migration after the occurrence of the first flood.

Úvod

Tato disertační práce se skládá ze čtyř kapitol, které jsem napsal během svého doktorského studia na CERGE-EI. Jako migrant jsem cítil, že moje osobní zkušenost s migrací má v sobě něco jedinečného, co bych měl sdílet, a tak tedy vzniklo téma této práce.

Všechny čtyři kapitoly analyzují téma pracovní migrace z různých perspektiv. V prvních dvou kapitolách zkoumám přistěhovaleckou politiku. V dynamickém světě, kde nové technologie rychle snižují náklady na mobilitu, se imigrační politika stává důležitým nástrojem při kontrole imigrace. Ve zbývajících dvou kapitolách se soustřeďuji na problematiku selekce do emigrace na příkladu Ukrajiny a na mobilitu uvnitř země na příkladu České republiky. Tyto jevy jsou důležité, protože určují směr a velikost změny blahobytu způsobeného mobilitou pracovních sil.

V první kapitole je vytvořen teoretický model, který vysvětluje, proč a kdy stát zavede vstupní omezení na počet a druh dovedností zahraničních pracovníků. Zavedením imigrační kvóty cílová země nepřímo ovlivňuje blahobyt země, odkud migranti pocházejí. Za určitých podmínek kvóta pozitivně ovlivňuje blahobyt země původu, protože snižuje rozsah negativního vlivu nových migrantů na mzdy v cílové zemi. V této kapitole dále popisuji, jak se kvóta mění v případě, že dvě země vytvoří politickou unii.

Ve druhé kapitole je vytvořen index imigrační politiky, který je ukazatelem laxnosti této politiky. Tento index má několik výhod oproti již existujícím ukazatelům. Je definován srovnatelně pro všechny země na světě a mění se v rámci párů cílové země a země původu i v čase. Když je tento index použit při odhadu, vysvětluje značnou část počtu migrantů jako stavové veličiny.. Také umožňuje vysvětlit složení pracovní síly migrantů z hlediska pohlaví a vzdělání.

Ve třetí kapitole je zkoumán způsob selekce migrantů z mé vlasti, Ukrajiny. Provedli jsme rozsáhlé šetření, ve kterém byly shromážděny informace o pozorovatelných charakteristikách migrantů a jejich výsledků na trhu práce před a po emigraci. S využitím této sady dat jsme zjistili, že ukrajinští migranti jsou pozitivně selektováni z hlediska věku, vzdělání a příjmu, kterého dosahovali před migrací. To se nicméně neodráží v jejich výsledcích na trhu práce, protože mnoho z nich je zaměstnáno v profesích pod svou nahlášenou úrovní vzdělání. To může být vysvětleno jako důsledek imigračních politik, vysokých nákladů na hledání, špatné přenositelnosti lidského kapitálu získaného na Ukrajině nebo individuálních nepozorovatelných schopností.

Ve čtvrté kapitole je zkoumána selekce do vnitřní migrace na východě České republiky. Tato část země je neustále předmětem relativně vysokých povodňových rizik z okolních vodních zdrojů. K pokrytí ztrát souvisejících s povodněmi a snížení zranitelnosti domácností začíná mnoho lidí dojíždět za prací do okolí větších měst, které nabízejí lepší pracovní příležitosti a vyšší plat. Zajímavé je, že zkoumaná oblast je charakterizována vysokým stupněm permanentní migrace z oblasti po výskytu prvních povodní.

Contents

1 Welfare Effects of Labor Migration under Policy Coordination **13**

1.1 Introduction 14

1.2 Documentation on migration 18

1.3 The Model 22

1.3.1 The setup 23

1.3.2 Free migration 24

1.3.3 Country A preference 26

1.3.4 Country B preference 29

1.3.5 Political union preference 31

1.3.6 Comparison of outcomes 33

1.3.7 Comparison with brain drain 37

1.4 Conclusions 38

A.1 Key migration statistics 43

2 Immigration Policy Index **47**

2.1 Introduction 48

2.2 The model of the determinants of migration 50

2.2.1 Version one: additive migration costs 51

2.2.2 Version two: multiplicative migration costs 52

2.3 Policy index design 53

2.4 Policy quasi-experiment setup 56

2.5 Data description 59

2.5.1 Construction of variables 59

2.5.2 Analysis of means of stocks 61

2.6 Econometric model and identification 63
2.7 Estimation results 65
2.7.1 Baseline regression 65
2.7.2 Nonlinear effects 67
2.7.3 Difference-in-difference estimates 70
2.7.4 Robustness check 72
2.7.5 Discussion of endogeneity 74
2.7.6 Discussion of results 76
2.8 Conclusion 77
B.1 Derivation of Gini Index 81
B.2 Tables 83
B.3 Figures 94

3 Migration from Ukraine: Brawn or Brains? New Survey Evidence 97

3.1 Introduction 98
3.2 Literature review 99
3.3 Survey design 102
3.4 Description of collected data 103
3.4.1 Descriptive statistics 103
3.4.2 Labor market outcomes: defining downshifters 108
3.4.3 Network effects 111
3.5 Econometric model and estimation 113
3.5.1 The model and identification 113
3.5.2 Estimation results 116
3.5.3 Robustness check 118
3.6 Conclusion 119
C.1 Tables 125
C.2 Figures 132

4 Commuting Patterns of Czech Households Exposed to Flood Risk from the River Bečva 135

4.1 Introduction 136

4.2 Survey design 138
4.3 Descriptive statistics 139
4.4 Estimation 142
4.4.1 Mincerian wage regression 142
4.4.2 Determinants of commuting 144
4.4.3 Commuting distance 149
4.5 Conclusion 151
D.1 Map of surveyed area 155
D.2 Definitions of variables 156

Chapter 1

Welfare Effects of Labor Migration under Policy Coordination

Abstract

The developed theoretical model analyzes the welfare effects of labor migration. I find that for the receiving country immigration enhances welfare as long as the marginal benefits to the locals' income exceed the social costs of immigration. Over-emigration of workers generated by free mobility is welfare detrimental to the source country because of the diaspora effect – migrants negatively affect their own income. The source country prefers to coordinate the immigration quota with the destination country because the coordinated solution internalizes the negative diaspora effect. Contrary to popular opinion, under certain conditions unilateral enforcement of the immigration quota also benefits the source country because it reduces the extent of the migrants' income decline.

Keywords: migration costs; wage effect; immigration policy; coordination
JEL Classification: F22; J15; D61; E61

An earlier version of the paper was published as CERGE-EI working paper No. 491. I am grateful to Byeongju Jeong for guidance in writing this paper. This research was supported by a grant from the CERGE-EI Foundation under a program of the Global Development Network. All opinions expressed are those of the author and have not been endorsed by CERGE-EI or the GDN.

1.1 Introduction

Zimmermann (1995) notes that the European countries[1] started restricting immigration flows after the first oil price shock in 1973, because it sparked fears of social tension and unemployment. Since then policy makers have been contemplating the optimal immigration policy, which, in principle, should address the issues of immigration quota, immigrants' characteristics, their rights to employment, family reunification, access to welfare and citizenship. In 2011 the UK introduced an annual immigration cap of 20 thousand on non-EU immigration plus 1 thousand in "exceptional talent" visas. Inter-company transfers, though not affected by this regulation, face restriction on earnings and duration of stay. In the US for the 2014 fiscal year the annual cap on H-1B category visas is 65 thousand. The intention is, due to high demand from employers, to distribute visas on a lottery basis. The Australian migration program for 2012–2013 is set for 190 thousand places, out of which 68% is reserved for skilled migrants. Following its immigration levels plan for 2013, Canada should accept 260 thousand migrants, out of which 62.3% are economic migrants.[2]

Despite the relevance of the issue for policy makers, the existing literature on the immigration policy is quite inconclusive. Giordani and Ruta (2011) note that existing theoretical models predict polarized immigration outcomes; either too many migrants or the closed door policy. The mismatch between theoretical predictions and practical outcomes calls for more research into the welfare effects of immigration, which is the driving force behind immigration policy.[3] Since the migration event involves three actors, namely, migrants, sending and receiving countries, there is a need for a theoretical framework that allows for a consistent comparison across them. This paper develops such a model. Specific questions addressed are: "How does immigration affect the welfare of the destination country?", "How does emigration affect the welfare of the source country?", "How many

[1]The reference is made to members of the European Economic Community as of 1974: France, Germany, Italy, Belgium, the Netherlands, Luxembourg, the UK, Ireland and Denmark.

[2]This information is taken from the respective official government web sites: `www.ukba.homeoffice.gov.uk`, `www.uscis.gov`, `www.cic.gc.ca`, and `www.immi.gov.au`.

[3]There is also insufficient empirical research on quantifying the immigration policy. Ortega and Peri (2009) create an index that measures the toughness of entrance and asylum laws. However, their measure is not heterogeneous across sending countries. Docquier et al. (2012) quantify the immigration policy by the fraction of refugees and females among migrants and existence of bilateral guest worker programs.

migrants move under free mobility?", and "What are the gains from coordinating a joint immigration policy?". Answering these questions in a unified framework sheds light on the structure of incentives of the actors, which helps explain the South-North migration.

This paper models only labor migration. According to data in Section 2 it accounts for no more than 40% of the incoming flows, the rest being mainly family-related migration. In most cases a worker moves first and is then followed by a spouse who is a "tied" mover (collective theory of family migration, see Mincer, 1978 and Rabe, 2011).

In his seminal paper Borjas (1995a) finds that immigration decreases natives' wages, redistributes wealth from workers to capital owners and creates an immigration surplus. The author argues that skilled immigration generates a larger surplus because skilled wages are more responsive to a shift in labor supply. For the US economy Storesletten (2000) finds that high-skilled migrants aged 40–44 are the most beneficial from the fiscal standpoint. For Germany, Akin (2012) finds that, at 2011 immigration rates the country's welfare is enhanced by around 3%. In his model with agents heterogeneous in wealth holdings Benhabib (1996) finds that if migrants decrease the capital/labor ratio, then those locals with above-average capital will have higher post-immigration income. The mirror image of this finding is also true, in that if the capital/labor ratio is increased by migrants, then those with below-average capital will have higher post-immigration income. Under majority voting the voters will be divided into those who prefer admitting migrants with either high or low wealth holdings. Bertoli and Brücker (2011) find that the shift towards a more selective immigration policy, without increasing the immigration volume, is always welfare detrimental to the source country. In their theoretical model, Razin and Sadka (1999) find that unskilled immigration into a welfare state with a pay-as-you-go pension system is strictly beneficial to all age groups. Fuest and Thum (2000) investigate the welfare effects of immigration when some sectors are unionized. They find that immigration is beneficial if the wage elasticity of labor demand in the competitive sectors is smaller than in the unionized one. In the opposite case small (large) scale immigration reduces (increases) locals' welfare.

Recent studies emphasize the role of social immigration costs, which include, but are not limited to, migrants' participation in welfare programs,[4] costs of border control and policing[5] and locals' dissatisfaction from having migrants in the neighborhood.[6] Giordani and Ruta (2011) define social costs as the fiscal and integration costs of the immigrant community. The authors argue for the presence of "congestion effects," when it becomes more difficult to integrate an additional migrant beyond a certain threshold. Schiff (2002) finds that immigrants decrease the social capital in the host society by increasing its diversity.

For the sending country the welfare effects of emigration are associated with how the income of stayers and migrants is affected by the marginal moving worker. I find that emigration decreases output in the source country by the migrants' wage leaving the income of stayers unaffected. Besides that, emigration generates a negative diaspora effect; a marginal migrant decreases the income of other migrants. The literature has two hypotheses on this issue; brain gain and brain drain. The brain drain literature finds that the emigration of skilled workers deprives the source country of the human capital that is important for its economic growth. Bhagwati and Hamada (1982) suggest taxing skilled emigrants and using the proceeds for developmental spending in the country of origin.[7] Burda and Wyplosz (1992) find that the market delivers too much migration relative to the social planner's outcome because it ignores the external social costs. The authors suggest introducing a labor subsidy in the source country and a one-shot tax on emigrants.

Cellini (2007) finds that emigration always lowers the welfare of the sending country because it decreases the average level of human capital and the labor productivity. The author argues that when immigration entails positive welfare effects for the receiving country, it is willing to accept more migrants than is optimal. The literature on brain gain (Mountford, 1997; Stark et al., 2004; Batista et al., 2012) concludes that under

[4]Ostrovsky (2012), Borjas (2011) provide evidence that migrants do participate in welfare programs. The rate and character of participation differs by the destination country, migrants' demographic characteristics and their duration of stay.

[5]The total budget of Frontex, the EU agency for border control, was EUR 86.8 million in 2011 (Frontex, 2011). This excludes the costs of policing measures of individual EU member states.

[6]Filer (1992) finds that the attractiveness of a city for the local workers negatively correlates with the volume of the recent immigrant population.

[7]Such a tax has not been introduced in practice, though some developing countries issue diaspora bonds, which bear a rather voluntary character (Ketkar and Ratha, 2010).

certain conditions the sending country benefits from skilled emigration because out of all prospective migrants who study more to boost their chances to emigrate, only a fraction will eventually emigrate and the remaining non-migrants will increase the average level of human capital. For example, Batista et al. (2012) find that for Cape Verde "a 10 pp increase in the probability of their own future migration improves the probability of completing intermediate secondary schooling by nearly 4 pp for individuals who do not migrate before age 16."

The model developed here predicts that a moving worker lowers the wage in the migrant sector of the destination country affecting the income of native and migrant workers. The negative effect of migrants on their own income (diaspora effect) is supported by several empirical studies. Using the example of the construction sector in Norway, Bratsberg and Raaum (2012) find that the wage effect varies across education groups. For the low- and medium-educated natives and migrants it is similar in magnitude and significance, whereas it is zero for the skilled natives and negative for the skilled migrants. A 10% increase in immigrant employment decreases native wages in construction by 0.6%. Borjas (1987b) finds that a 10% increase in the supply of immigrants reduces the immigrant wage by about 10%. LaLonde and Topel (1991) report that a 10% increase in new immigration reduced wages of new immigrants by 0.24%. The studies find that in the long run the reported negative effect disappears because of the adjustments: locals out-migrate from areas (Filer, 1992) or exit sectors (Bratsberg and Raaum, 2012) with a high concentration of immigrants, and new industries locate in places with a relatively large supply of unskilled labor.

In the paper I first provide documentation on migration and evidence of coordination of immigration policies on the EU level. In the model section I first describe economies of countries A and B. Then I formalize the migration preference of individual workers, the immigration preference of receiving Country A, the emigration preference of sending Country B and the preference of the political union. I further compare the four outcomes and conclude.

1.2 Documentation on migration

Migration is a bilateral phenomenon. It is established between a pair, or groups, of countries and evolves over time. The world migration picture is quite diverse and dynamic. Data in Figure 1.1 suggest that in 1990 migration within the developing world (South–South) ranked first in volume and accounted for almost 40% of the total stock of migrants,[8] whereas migration within the developed world (North–North) and the developing to developed world (South–North) ranked second and third with respective shares of 27.1% and 25.7%. By 2000 the world workforce had become significantly more mobile and total migration grew by around 38.2%. The growth was primarily driven by the increase in South-North migration (86.2%), which surpassed all other flows in volume and in 2000 totaled 74.3 million people or 34.7% of the total migrant stock.

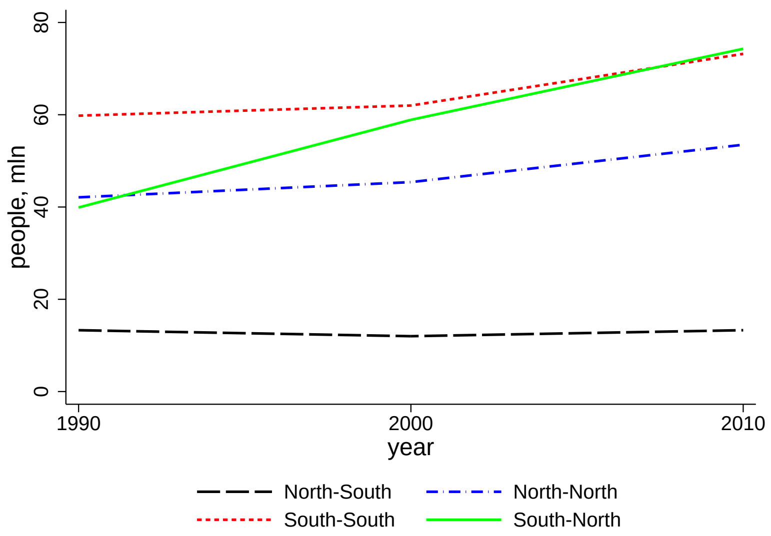


Figure 1.1: *International migrant stock (in mln) by source and destination region.* *Source:* UN (2011).

Several factors stand behind the rapid growth of South–North migration. The definition of "North" in 2000 includes more countries than in 1990. The developed economies need young migrant workers to satisfy labor shortages and support the ageing population,

[8]There is no convention on what defines a migrant and destination countries use their national definition. To avoid data inconsistency in the cross-country comparison, OECD standardizes the migration statistics. In many instances OECD and UN define a migrant as a foreign-born individual. Unless otherwise noted, I will keep to this definition throughout the text. See OECD (2012) and UN (2011) for a detailed discussion on national definitions.

among other reasons. The developing world has got wealthier and migration costs have declined in many ways, making it easier for people in the source countries to satisfy the migration budget constraint.

Realization of the event of economic migration entails two selection effects. Only those individuals (families) emigrate who expect to benefit from emigration.[9] Out of the pool of potential migrants the immigration policy admits those who meet selection criteria as long as the quota has not been exhausted.[10] These two types of selection affect all aspects of migration, viz. volume of migrants, their demographic characteristics and details of economic activity. Since only the migration outcome is observed, it is now being actively discussed how to identify the contribution of each selection type.

Table A.1.1 contains basic standardized descriptive statistics. Comparing inflows and outflows in 2010 most OECD countries, except Ireland and Greece, are the net recipients of migrants. For large receiving countries in per capita terms, such as Norway, Switzerland and Austria, more than two thirds of migrants come from within the European Union, which reduces the demand of these countries for foreign labor from outside the EU. In Italy and the UK, the share of labor recruitment from outside the EU is 40.5% and 33.3% of total inflows respectively, whereas in most other countries this share rarely exceeds 20%. Family migration accounts for a significant proportion in almost all destination countries, the largest being in the US (66.3%), France (42.9%) and Sweden (39.6%). The Scandinavian countries are active in the humanitarian mission: 18.7% in Sweden, 17.4% in Finland and 9.5% of inflows in Norway are humanitarian migrants.

The migration flows translate into stocks via the law of motion. In some countries with relatively high inflows in 2010, the stocks are also high, which suggests that immigration is persistent and of a more permanent type. For example, in Switzerland, Sweden and Austria the stock of foreign born people is 26.6%, 14.8% and 15.7% of the local population respectively. On the contrary, Ireland and the US had relatively low inflows in 2010 (3.9 and 3.4 migrants per one thousand of local residents), but the stocks are relatively high: 17.3% and 12.2% of the local population respectively, which suggests that the inflows

[9]Borjas (1987a) and Clark et al. (2007) are the key studies in the migration literature, whereas Heckman (1979) develops a general approach to address the sample selection.

[10]In 2012 the refusal rate in Canada was 22.5% for permanent residence and 15.8% for temporary residence applications. For comparison, in the US in FY 2012 the refusal rate for non-immigrant visas was 19.6%.

slowed down prior to 2010. The stock of foreign nationals is usually smaller than the stock of those foreign born, because migrants naturalize over time and disappear from the statistics on foreign nationals.

In all countries for which data are available, except Hungary and the US, the unemployment rate among the foreign born exceeds the unemployment rate of the native born. Two comments are relevant here. Firstly, it is an established fact in the literature that migrants are disadvantaged in the labor market for some time after their arrival (literature on assimilation). Secondly, the foreign born might be different in some underlying characteristics (education, unobservable skills and talent) for which they get penalized in the labor market. The data suggest that for all countries in the sample a migrant is more likely than a native person to have less than upper secondary education. At the same time in Austria, Hungary, Switzerland, Germany, Luxembourg and Sweden migrants are also more likely than the locals to have tertiary education. This observation suggests the polarization of migrants' education (skills); a migrant is likely to be either in the low or high education category.

The evidence on migrants' educational attainments should be reflected in their employment details. Table A.1.2 illustrates data on employment sector and occupations of the foreign born. In the countries considered, with the exception of Greece and Italy, the share of migrants employed in the service sector exceeds 20%. In the Czech Republic and Germany more than one quarter of migrants are employed in mining, manufacturing and energy and for other countries, with the exception of Luxembourg, this share is above 10%. In all countries considered 10–15% of migrants are employed in trade. In Norway, Sweden and Denmark around 20%, in the Netherlands, Ireland and the UK around 15% of migrants are employed in the health sector. In contrast, migrants are highly unlikely to be employed in the agriculture and fishing, household (except Italy and Greece), education and administrative sectors.

The data in Table A.1.2 also suggest that migrants are less likely to be employed in more skill demanding occupations. In elementary occupations migrants are overrepresented compared to local workers in all countries considered. In the professionals, senior officials and managers category migrants are more likely than locals to be employed only in Austria, Hungary, Switzerland and Luxembourg.

The evidence thus suggests that some sectors (services, trade, mining and manufacturing, less so health care and household) and some occupations (elementary occupations, less so clerks and skilled trades) are more prone to employing migrants than other sectors and occupations. The nature of this observation is driven by immigration policy, migrants' educational attainments, language barriers, poor cross-border transferability of skills (Mattoo et al., 2008) as well as licensing and certification requirements (particularly in health care). The immigration policies of major receiving destinations often favor brain over brawn. Skill selective immigration policies have been adopted in the European Union, UK, Canada and Australia.[11] Besides that, the EU member countries coordinate the immigration policies because of the common labor market within the European Economic Area.

Despite the absence of a unified immigration system on the EU level, significant progress has been made in harmonizing rules regarding the admission and treatment of migrants. There is a clear trend in favoring skilled immigration (EU Blue Card Directive 2009/50/EC and Directive 2005/71/EC). These directives stipulate simplified visa and admission procedures for the respective categories and a fast track to permanent residence upon satisfaction of certain criteria. Legal long-term migrants have the right to bring in their families, obtain access to health care, education and public services (Directive 2003/86/EC, Directive 2003/109/EC) and there are common rules for the admission of students (Directive 2004/114/EC). A significant achievement is the agreement on the single residence permit that stipulates the issue of a single document that encompasses the residence and work permit (Directive 2011/98/EU).

The external dimension of the EU immigration policy includes active cooperation with countries of origin and transit in terms of tighter border enforcement and control, cooperation on data sharing and readmission of undocumented migrants. The Global Approach to Migration set out in the Stockholm program for 2010–2014 calls for actions that ensure efficient management of migration flows to benefit all countries concerned. Three types of agreement with non-EU (third) countries are actively being used: mobility partnerships, readmission agreements and visa facilitation agreements.

[11] For detailed description see OECD (2013) for the EU, Mavroudi and Warren (2013) for the UK, Gera and Songsakul (2007) for Canada and Miller (1999) for Australia.

The mobility partnerships aim at better management of immigration flows via development programs in the migrant source country and circular mobility programs. The intention here is to make a difference in the country of origin, before the person actually becomes a migrant. The projects implemented within each partnership depend on the needs of a particular country, though there is a preference to encourage legal temporary migration, better border control, and information sharing to discourage potential undocumented migration. Mobility partnerships work on a "more-for-more" principle, when more cooperative third countries get less restrictive visa regimes. As of the end of 2012 mobility partnerships were signed with four countries: Georgia, Moldova, Armenia and Cape Verde.

The readmission agreements are aimed at combatting illegal migration. They establish a procedure under which the source country accepts undocumented migrants who either originate from that country or used it as a transit country. Despite the fact that only half of the repatriated cases end up in readmission, Billet (2010) argues that the readmission agreements are a milestone in coordination of the immigration flows between the EU and large sending countries. In exchange for the cooperation on readmissions the EU may grant visa facilitation agreements that simplify visa requirements for seasonal and temporary migrants from cooperating third countries. The readmission and visa facilitation agreements have been signed with Moldova, Georgia, Ukraine and Russia, amongst other countries.

1.3 The Model

The world consists of two regions: North and South. North represents developed migrant receiving countries, for example OECD members, and South represents developing migrant sending countries, for example republics of the Former Soviet Union, India or Latin America.[12] Country A is an average country of North and Country B is a large representative country of South. Migration statistics presented in Table A.1.2 suggests division of the economies of both countries into migrant and non-migrant sectors. In Country A the migrant sector employs native and migrant workers and can be thought

[12]See UN (2011) and Marchiori et al. (2013) for a more extensive definition of North and South.

of as elementary, clerks and service occupation in manufacturing, trade or health care. The non-migrant sector employs only native workers. In Country B workers can emigrate only from the migrant sector. In either country assignment to sectors is exogenous and workers cannot switch sectors. Each worker inelastically supplies one unit of labor.

1.3.1 The setup

Country A produces the final good competitively with the Cobb-Douglas constant returns to scale technology:

$$Y^A = Z^A \left(L^A\right)^{\beta} \left(H^A + M\right)^{1-\beta},$$

where Z^A is the total factor productivity, H^A and L^A is native labor employed in migrant and non-migrant sectors respectively, M is migrant labor. Let N^A be the total native population.

Under the assumption of competitive factor markets the wage in each sector equals the marginal product of its workers:

$$w_L^A = \beta Z^A \left(L^A\right)^{\beta-1} \left(H^A + M\right)^{1-\beta} = \beta \frac{Y^A}{L^A}, \tag{1.1}$$

$$w_H^A = (1-\beta) Z^A \left(L^A\right)^{\beta} \left(H^A + M\right)^{-\beta} = (1-\beta) \frac{Y^A}{H^A + M}. \tag{1.2}$$

The output is divided between the migrant and non-migrant sectors in shares $(1-\beta)$ and β.

Migrant workers come from less developed Country B with total population N^B. Output in Country B is produced competitively according to the Cobb-Douglas technology with constant returns to scale:

$$Y^B = Z^B \left(L^B\right)^{\gamma} \left(H^B - M\right)^{1-\gamma},$$

where H^B is labor employed in the migrant sector, M are emigrants, L^B are workers employed in the non-migrant sector who cannot emigrate.

Similarly, under the assumption of competitive factor markets the wages in Country

B are:

$$w_L^B = \gamma Z^B (L^B)^{\gamma-1}(H^\beta - M)^{1-\gamma} = \gamma \frac{Y^B}{L^B}, \tag{1.3}$$

$$w_H^B = (1-\gamma) Z^B (L^\beta)^{\gamma}(H^\beta - M)^{-\gamma} = (1-\gamma) \frac{Y^B}{H^B - M}. \tag{1.4}$$

The output is divided between the migrant and non-migrant sectors in shares $(1-\gamma)$ and γ. To generate individual migration incentives I assume that Country A is technologically more advanced than Country B.

1.3.2 Free migration

Each worker employed in the migrant sector of Country B faces the choice whether to stay and get wage w_H^B for the unit of labor supplied, or emigrate to Country A and get w_H^A, $w_H^A > w_H^B$. In order to emigrate worker i must pay $c(i)$ for the migration costs.[13] Worker's maximization problem is formalized as follows:

$$\begin{aligned} \max \quad & \{w_H^B,\ w_H^A - c(i)\} \\ \text{s.t.} \quad & \text{equations (1.2) and (1.4).} \end{aligned} \tag{1.5}$$

The worker emigrates if the wage gain exceeds or equals the individual migration costs and stays otherwise. The individual index i ranks workers according to their migration costs; higher values of the index corresponds to higher costs, $i \in [0, H^B]$ (see Figure 1.2). Workers with low costs emigrate first. Worker with $i = 0$ has zero migration costs and gains $w_H^A - w_H^B$ from emigration. The marginal worker's costs increase by $\frac{\bar{C}}{H^B}$ and worker i gains $w_H^A - w_H^B - i\frac{\bar{C}}{H^B}$ from emigration. This is equivalent to saying that $c(i) \sim U[0, \bar{C}]$.[14] Let M^M denote the market level of emigration, which is determined from the following equations:

$$\breve{w}_H^A - \breve{w}_H^B = M^M \frac{\bar{C}}{H^B}, \tag{1.6}$$

where $\breve{w}_H^A$ and $\breve{w}_H^B$ are wages at M^M.

[13] In broader migration literature the individual migration costs include material costs of the move, costs of social exclusion and discrimination. Carrington et al. (1996), Beine et al. (2011) find that the migration costs decline as the stock of migrants of the same nationality grows.

[14] The distribution assumption does not affect the model result, although it makes it more trackable.

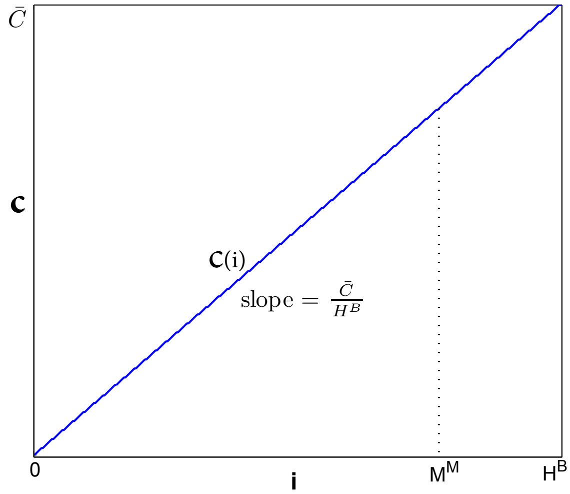


Figure 1.2: *Visualization of individual migration costs.*

Given the migration level M^M and the assumption of the uniform distribution of the costs, the total migration costs paid, which is the triangular area below the diagonal line in Figure 1.2, equal $\frac{\left(M^M\right)^2 \bar{C}}{2H^B}$.

If Country A becomes relatively more technologically developed, *ceteris paribus*, M^M will increase. If Country A employs more people in the migrant sector, the wage in that sector declines thus reducing M^M. If $\bar{C}$ increases, the average migration costs increase, thus resulting in less migration. M^M does not depend on the total population in both countries. However, it does depend on the distribution of workers across the two sectors.

It must further be noted, that the individual decision rule in equation (1.5) accounts for the fact that in the migrant sector a moving worker decreases the wage in Country A and increases the wage in Country B. This, however, ignores a number of the welfare effects, for example, a change in income of native and migrant workers induced by the change in wages (Card, 1990; Bratsberg and Raaum, 2012) as well as the social costs incurred from immigration (Giordani and Ruta, 2011).

1.3.3 Country A preference

Country A maximizes the welfare of its native workers by choosing the volume of migrants M to accept for employment in the migrant sector. The maximization problem is defined as follows:

$$\begin{aligned} \max_{\{M\}} W^A(M) &= L^A w_L^A + H^A w_H^A - \frac{A}{2}\left(\frac{M}{N^A}\right)^2 N^A \\ \text{s.t.} &\quad \text{equations (1.1) and (1.2)} \\ &\quad M \geq 0. \end{aligned} \tag{1.7}$$

The first two terms of the welfare function is the income that accrues to the native workers minus wages paid to migrants. The third term is the social immigration costs incurred by the receiving country.[15] This term expresses in monetary terms the value of all costs that the country incurs from accepting migrant workers: border controls and policing, integration and language courses or simply the natives' dissatisfaction from having migrants around.

The welfare effect of immigration is derived by differentiating (1.7) w.r.t. M. After rearrangement I obtain:

$$\frac{\partial W^A}{\partial M} = L^A \frac{\partial w_L^A}{\partial M} + H^A \frac{\partial w_H^A}{\partial M} - \frac{A}{N^A} M = \frac{\beta w_H^A}{N^A}\left(1 - \frac{H^A}{H^A + M}\right) - \frac{A}{N^A} M.$$

Since $\frac{\beta w_H^A}{N^A}\left(1 - \frac{H^A}{H^A+M}\right) > 0$ for $M > 0$, the natives' income is strictly increasing in the number of migrants. Disregarding the social costs, the native workers are strictly better off from the marginal migrant. To see why it is so, consider the migrant's effect on output:

$$w_H^A = \frac{\partial Y^A}{\partial M} = \frac{\partial}{\partial M}\left(L^A w_L^A + \left(H^A + M\right) w_H^A\right) = L^A \frac{\partial w_L^A}{\partial M} + H^A \frac{\partial w_H^A}{\partial M} + M \frac{\partial w_H^A}{\partial M} + w_H^A.$$

A migrant is paid his marginal product and his arrival generates two more effects which cancel out: a positive effect on wage in the non-migrant sector and a negative

[15] The welfare function disregards the welfare of foreign workers. This is a standard assumption in most cited studies. For example, Giordani and Ruta (2011) use a different functional form, but the function properties remain the same.

effect on wage in the migrant sector.

$$\underbrace{L^A\frac{\partial w_L^A}{\partial M}+H^A\frac{\partial w_H^A}{\partial M}}_{\text{effect on natives, } > 0}+\underbrace{M\frac{\partial w_H^A}{\partial M}}_{\substack{\text{effect on}\\ \text{diaspora, } \leq 0}}=0. \tag{1.8}$$

The negative effect on the wage in the migrant sector reduces the income of the native and migrant workers. Reduction of the natives' income is smaller in absolute value than the increase in the non-migrant sector. For this reason, disregarding the social costs, immigration always increases the natives' income. The positive effect on the locals is referred to in the literature as the "immigration surplus" (Borjas, 1995b; Giordani and Ruta, 2011).

The third term in equation (1.8) is the effect of migrants on their own income, which I call the diaspora effect. This effect is defined to be:

$$M\frac{\partial w_H^A}{\partial M}\begin{cases} = 0 & \text{if } M = 0, \\ < 0 & \text{if } M > 0. \end{cases}$$

The diaspora effect does not affect the welfare of Country A, because the migrants take away the negative effect on themselves. The diaspora effect in a crucial way affects the welfare of Country B, which is considered in detail in Section 3.4.

It costs $\frac{A}{N^A}$ in social costs to accept the marginal migrant. The welfare effect of immigration is:

$$\frac{\partial W^A}{\partial M}\begin{cases} = 0 & \text{if } M = 0, \\ \geq 0 & \text{if } \frac{\beta w_H^A}{N^A}\left(1-\frac{H^A}{H^A+M}\right) \geq \frac{A}{N^A}M, \\ < 0 & \text{if } \frac{\beta w_H^A}{N^A}\left(1-\frac{H^A}{H^A+M}\right) < \frac{A}{N^A}M. \end{cases}$$

When there are no migrants in Country A, $M = 0$, the first migrant does not generate the diaspora effect, therefore the effect on the locals is zero. When migration continues, the marginal migrant positively affects the natives' income, and negatively affects the income of migrants already in the country through the diaspora effect. Country A will continue to accept migrants until the marginal increase in the natives' income equalizes the marginal increase in the social costs. The last migrant allowed in increases the locals' income by strictly as much as he increases the social costs.

Denote the optimal volume of migrants that solves maximization problem (1.7) by M^A. Using the welfare effects as the sole determinant of the immigration policy, two immigration policy profiles of Country A are considered:

1. Immigration quota:

$$M^A = \left[\frac{N^A Z^A \beta(1-\beta)(L^A)^\beta}{A}\right]^{\frac{1}{1+\beta}} - H^A, \tag{1.9}$$

2. Immigration ban:

$$M^A = 0.$$

The immigration quota defines the immigration volume that maximizes the welfare of Country A. It can happen that the quota is not exhausted, in which case the first-best outcome is not achieved. The country will not accept more than the quota, because of the social costs. The quota is strictly increasing in the total workforce, N^A, total factor productivity, Z^A, and decreasing in the social cost parameter A. When the native workforce is predominantly employed in the non-migrant sector, the country accepts many migrants, because the welfare can be increased by extending employment in the migrant sector. The immigration quota M^A is concave in β. For low values of β the migrant sector is more important in production and has a high marginal effect on the welfare. As β increases the marginal effect on the quota is positive until a certain point, after which it becomes negative. For small and large values of β the quota is smaller than for intermediate values.

When migrants do not cause any social costs, i.e. $A = 0$, the country accepts infinitely many migrants because the marginal effect on the natives' income is strictly positive, as derived in equation (1.8). Similar "open door" immigration policy predictions are confirmed by Bianchi (2013) and Giordani and Ruta (2011).

1.3.4 Country B preference

Migrant workers come to Country A from a less developed Country B. If Country B could choose how many emigrants to send, it would do so by maximizing the welfare of its emigrants and stayers. It thus solves the following maximization problem:

$$\begin{aligned} \max_{\{M\}} W^B(M) &= M w_H^A - M\frac{M\bar{C}}{2H^B} + L^B w_L^B + \left(H^B - M\right) w_H^B \\ \text{s.t.} \quad & \text{equations (1.1), (1.2), (1.3) and (1.4)} \\ & M \geq 0. \end{aligned} \tag{1.10}$$

The first term of the welfare function is the income of migrants, the second term is the total individual migration costs, the third and fourth terms are the income of stayers in Country B. To learn the welfare effects of emigration I have to differentiate (1.10) w.r.t. M. After rearrangement I obtain:

$$\frac{\partial W^B}{\partial M} = \underbrace{w_H^A - w_H^B - \frac{M\bar{C}}{H^B}}_{\text{net gain from emigration}} + \underbrace{M\frac{\partial w_H^A}{\partial M}}_{\text{diaspora effect, } \leq 0} + \underbrace{L^B\frac{\partial w_L^B}{\partial M} + \left(H^B - M\right)\frac{\partial w_H^B}{\partial M}}_{\text{effect on stayers, } = 0}. \tag{1.11}$$

Starting from no emigration, $M = 0$, the first migrant gains w_H^A, loses w_H^B and pays nothing in migration costs. For the first migrant, the diaspora effect is zero, because no migrants in Country A are affected by the wage reduction. For $M > 0$ each marginal migrant will reduce the wage paid to the first migrant, thus generating the negative diaspora effect.

Emigration has two effects on stayers; increase of wage in the migrant sector and decrease of wage in the non-migrant sector. These two effects cancel out:

$$-w_H^B = \frac{\partial Y^B}{\partial M} = \frac{\partial}{\partial M}\left(L^B w_L^B + \left(H^B - M\right) w_H^B\right) = L^B\frac{\partial w_L^B}{\partial M} + \left(H^B - M\right)\frac{\partial w_H^B}{\partial M} - w_H^B.$$

The welfare effect of emigration on Country B is thus:

$$\frac{\partial W^B}{\partial M}\begin{cases} = w_H^A - w_H^B \text{ if } M = 0, \\ \geq 0 \text{ if } w_H^A - w_H^B \geq \frac{M\bar{C}}{H^B} - M\frac{\partial w_H^A}{\partial M} \text{ and } M > 0, \\ < 0 \text{ if } w_H^A - w_H^B < \frac{M\bar{C}}{H^B} - M\frac{\partial w_H^A}{\partial M} \text{ and } M > 0. \end{cases}$$

The first migrant increases welfare by exactly as much as his net private gain from emigration. From then on, each marginal migrant increases the welfare by less than the private gain from emigration because of the negative diaspora effect. As the number of migrants increases, the marginal welfare gain declines because the wage differential narrows, the individual migration costs increase and the diaspora effect grows. The country prefers to send migrants as long as the wage differential exceeds the marginal migration costs and the marginal decline in income of the diaspora. The wage gain for the last migrant that Country B wants to send exactly equals the marginal migration costs plus the marginal increase in the diaspora effect.

I use M^B to denote the emigration level that solves maximization problem (1.10) and $\widehat{w}_H^A$, $\widehat{w}_H^B$ to denote wages at M^B. Two emigration profiles of Country B are considered:

1. Emigration quota:

$$\widehat{w}_H^A - \widehat{w}_H^B = \frac{M^B\bar{C}}{H^B} + M^B\frac{\beta\widehat{w}_H^A}{H^A + M^B}, \tag{1.12}$$

2. Emigration ban:

$$M^B = 0.$$

Comparing equations (1.12) and (1.6) one can notice that they differ only by the term that captures the diaspora effect. This means that for $M > 0$ Country B always prefers to have fewer migrants than the volume that self-establishes under free migration.

The emigration quota depends on the sectoral distribution of workers. Larger employment in the migrant sector in County A (B) will reduce (increases) the wage differential, thus driving down (up) the emigration quota. If Z^A $\left(Z^B\right)$ increases, M^B will increase (decrease), because the difference in wages rises (falls). If the average migration costs decline, i.e. $\bar{C}$ falls, Country B prefers to have more migrants.

1.3.5 Political union preference

Suppose that the two countries form a political union. It is then interesting to know how many migrants the union would like to have. The union solves the following maximization problem:

$$\begin{aligned}
\max_{\{M\}} W^U(M) &= L^A w_L^A + H^A w_H^A - \frac{A}{2}\left(\frac{M}{N^A}\right)^2 N^A + M w_H^A - M\frac{M\bar{C}}{2H^B} + \qquad (1.13)\\
&+ L^B w_L^B + \left(H^B - M\right) w_H^B \\
\text{s.t.} \quad & \text{equations (1.1), (1.2), (1.3) and (1.4)} \\
& M \geq 0.
\end{aligned}$$

The first three terms of the objective function is the income that accrues to the natives of Country A, net wages paid to migrants and the social immigration costs. The fourth and fifth terms are the migrants' income minus the migration costs. The second line is the income of stayers in Country B.

To learn the welfare effects of migration I have to differentiate (1.13) w.r.t. M. After rearrangement I obtain:

$$\begin{aligned}
\frac{\partial W^U}{\partial M} = \quad & \underbrace{L^A\frac{\partial w_L^A}{\partial M} + H^A\frac{\partial w_H^A}{\partial M} + M\frac{\partial w_H^A}{\partial M}}_{=0} + \underbrace{w_H^A - w_H^B - M\frac{A}{N^A}}_{\text{net gain from emigration}} - \underbrace{M\frac{\bar{C}}{H^B}}_{\text{social cost}} + \\
& + \underbrace{L^B\frac{\partial w_L^B}{\partial M} + H^B\frac{\partial w_H^B}{\partial M} - M\frac{\partial w_H^B}{\partial M}}_{\text{effect on stayers' income, } = 0}.
\end{aligned}$$

The marginal migrant with index i gains w_H^A, looses w_H^B and pays $i\frac{\bar{C}_H}{H^B}$ for the immigration costs. For accepting this migrant the union pays $\frac{A}{N^A}$ in the form of social costs. The marginal net gain to the union is thus $w_H^A - w_H^B - i\frac{\bar{C}_H}{H^B} - \frac{A}{N^A}$. Since the union cares about the welfare of all its workers irrespective of their country profile, migrants stop being migrants and the pronounced diaspora effect is internalized as it is shown in equation (1.8).

The union prefers to have migrants as long as the wage gain from migration exceeds the marginal individual and social costs. For the last migrant the wage gain will exactly equal the marginal individual and social costs. The welfare effect of migration in the union is as follows:

$$\frac{\partial W^U}{\partial M} \begin{cases} = w_H^A - w_H^B - \frac{A}{N^A} \text{ if } M = 0, \\ \geq 0 \text{ if } w_H^A - w_H^B \geq \frac{M\bar{C}}{H^B} + \frac{A}{N^A}M \text{ and } M > 0, \\ < 0 \text{ if } w_H^A - w_H^B < \frac{M\bar{C}}{H^B} + \frac{A}{N^A}M \text{ and } M > 0. \end{cases}$$

Let M^U denote the optimal migration level within the union, $\overline{w}_H^A$ and $\overline{w}_H^B$ denote wages at M^U. Then two migration profiles of the union are considered:

1. Migration quota:

$$\overline{w}_H^A - \overline{w}_H^B = \frac{M^U \bar{C}_H}{H^B} + \frac{A}{N^A}M^U, \tag{1.14}$$

2. Migration ban:

$$M^U = 0.$$

The union migration policy is the weighted average of the individual migration profiles of the two countries. If Country A accepts migrants more aggressively than Country B wishes to send them, the optimal volume for the union will be below that of Country A and above than of Country B. In the opposite case, when Country A wishes to accept less migrants than Country B wants to send, the union preference will be above that of Country A and below that of Country B.

When the social costs of immigration are reduced to zero, $A = 0$, the union preference equals the free market outcome because the migrants in the union do not impose any negative effects on the income of other migrants. This intuition is formalized in Proposition 1.

1.3.6 Comparison of outcomes

In this section I compare the four migration outcomes: free market, preference of Country A, preference of Country B and preference of the political union. As a starting point, let me recall the optimal migration levels:

$$\text{Market:} \quad \breve{w}_H^A - \breve{w}_H^B = \frac{M^M \bar{C}}{H^B}, \tag{1.15}$$

$$\text{Country A:} \quad \beta \widetilde{w}_H^A \left(1 - \frac{H^A}{H^A + M^A}\right) = \frac{AM^A}{N^A}, \tag{1.16}$$

$$\text{Country B:} \quad \widehat{w}_H^A \left(1 - \frac{\beta M^B}{H^A + M^B}\right) - \widehat{w}_H^B = \frac{M^B \bar{C}}{H^B}, \tag{1.17}$$

$$\text{Union:} \quad \overline{w}_H^A - \overline{w}_H^B = \frac{M^U \bar{C}}{H^B} + \frac{AM^U}{N^A}, \tag{1.18}$$

where $\breve{w}_H^S$, $\widetilde{w}_H^S$, $\widehat{w}_H^S$ and $\overline{w}_H^S$, $S = A, B$, are wages in migrant sectors at respective migration levels. The four outcomes are depicted in Figure 1.3.

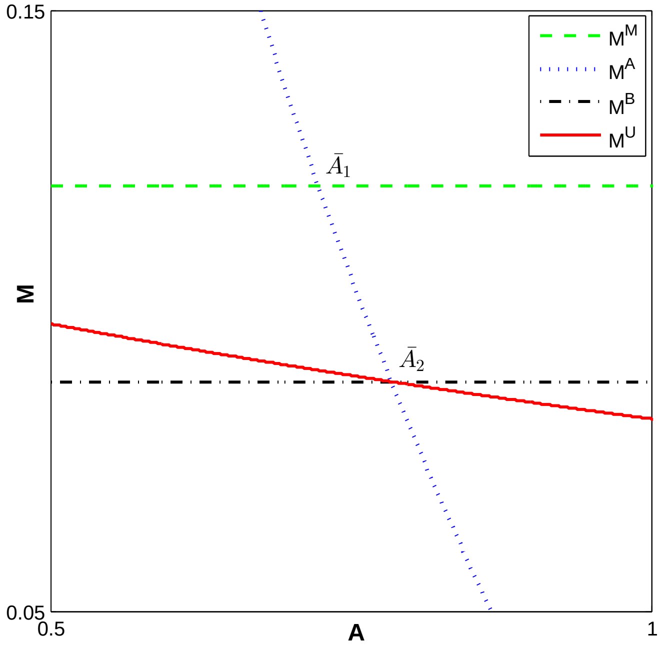


Figure 1.3: *Illustration of the migration outcomes for the following parameter values:* $Z^A = 2$, $Z^A = 1.5$, $N^A = N^B = 1$, $H^A = H^B = 0.5$, $\beta = \gamma = 0.5$, $\bar{C} = 0.15$.

The immigration quota M^A defines the volume of immigrants that is best for receiving Country A. When the social immigration cost parameter A declines, the country accepts more migrants.

The market outcome defines a migration level when workers are allowed to move freely. The worker's decision to move, as given by maximization problem (1.5), contrasts the wage differential and the marginal migration costs. It ignores the social immigration costs and the negative effect of migrants on their own income through the diaspora effect.

The preference of Country B defines the volume of migrants that is best for its welfare, which consists of the income of workers in the non-migrant sector, stayers in the migrant sector and the emigrants. When the wage differential is sufficiently high, Country B can increase its welfare by expatriating some of its workers to work in Country A where they are more productive. As the number of emigrants increases, the wage in the destination country falls, thus decreasing the income of migrants (diaspora effect). The optimal emigration quota for Country B is when the wage differential equals the marginal decrease in the migrants' income plus the marginal migration costs. This compares to the market migration level, which disregards the migrants' effect on their own income. Proposition 1 shows that the negative diaspora effect decreases the optimal volume of migrants for the source country relative to the free market level.

The union outcome describes the case when both countries can agree on such a level of migration, which is best for the world. The union quota is thus a weighted average of the preferences of sending and receiving countries; it accounts for the wage differential, social and individual costs as well as the negative diaspora effect. Since in the union migrant workers stop being migrants (their well-being is cared for by the planner) the negative diaspora effect is internalized. For this reason Country B always benefits from coordination. The internalization of the negative diaspora effect is a form of distribution effect. If the union level of migration is not achievable for some reason, then side payments can be used to achieve higher union welfare. For example, one could think of a model whereby the union outcome is characterized by migration restrictions and side payments from the destination to a source country (Stark et al., 2012).

Two propositions below rank the migration outcomes depending on the value of the social cost parameter.

Proposition 1. *If $A = 0$, the outcomes are ranked $M^A > M^M = M^U > M^B$.*

Proof. If $A = 0$, then from equations (1.9) or (1.16) it follows that $M^A \to \infty$. Next,

$M^M = M^U$ because equations (1.15) and (1.18) are the same. Further, I subtract equation (1.15) from (1.17) to obtain:

$$\widehat{w}_H^A \left(1 - \frac{\beta M^B}{H^A + M^B}\right) - \breve{w}_H^A + \breve{w}_H^B - \widehat{w}_H^B = \frac{\bar{C}}{H^B}\left(M^B - M^M\right). \tag{1.19}$$

If $M^B > M^M$, then the LHS of (1.19) is negative but the RHS is positive, which is a contradiction. If $M^B = M^U$, then the LHS is negative but the RHS $= 0$, which is again a contradiction. Then $M^B < M^M$ is the true relationship, because it does not produce a contradiction.

By the transitivity property the ranking $M^A > M^M = M^U > M^B$ follows. □

Proposition 1 establishes the first key result of the paper – over-emigration. If workers are allowed to move freely, the market outcome delivers more migrants than the quota of Country B, $M^M > M^B$. In other words, more people move than is optimal for the source country. Compared to findings of the brain drain literature, this result suggests that emigration of a marginal worker decreases output in the sending country by the worker's wage; emigration does not decrease the income of stayers; and, finally, the emigration of $M^M - M^B$ extra migrants is harmful to the source country because they excessively decrease the income of M^M migrants who are already in the destination country.

Further, since the union preference internalizes the negative diaspora effect and when the social immigration costs are zero, the union quota equals the free market level of migration. The political union prefers to have as many migrants as workers who wish to move. This result also follows from the application of the First Welfare Theorem. In the absence of social immigration costs the receiving country prefers an open door immigration policy, because the marginal benefit of an additional migrant is strictly positive. This result is not uncommon in the literature: Giordani and Ruta (2011) and Bianchi (2013) are most recent studies that confirm it.

Proposition 2. *There exist $\bar{A}_1$ and $\bar{A}_2$, such that:*

$$M^A > M^M > M^U > M^B \quad \textit{if} \quad A < \bar{A}_1, \tag{1.20}$$

$$M^M > M^A > M^U > M^B \quad \textit{if} \quad \bar{A}_1 < A < \bar{A}_2, \tag{1.21}$$

$$M^M > M^B > M^U > M^A \quad \textit{if} \quad A > \bar{A}_2. \tag{1.22}$$

Proof. The three cases are depicted in Figure 1.3. It follows from Proposition 1 that M^M and M^A do not depend on A, they are parallel lines with M^M above M^B, $\frac{\partial M^M}{\partial A} = \frac{\partial M^B}{\partial A} = 0$, $M^M > M^B$.

M^A is a continuous function decreasing in A. For small values of A M^A is above M^M and for large values M^A is below M^B. There exist a unique point $\bar{A}_1$ at which M^A intersects M^M, such that if $A < \bar{A}_1$, $M^A > M^M$, and conversely, if $A > \bar{A}_1$, $M^A < M^M$. Similarly, there exist a unique point $\bar{A}_2$ at which M^A intersects M^B, such that if $A < \bar{A}_2$, $M^A > M^B$, and, conversely, if $A > \bar{A}_2$, $M^A < M^B$. Since M^A is a downward sloping line, it first crosses M^M and then M^B, $\bar{A}_1 < \bar{A}_2$.

M^U is a continuous function decreasing in A. For $A = 0$ $M^U = M^M$ (Proposition 1), and for $A > 0$ $M^U < M^M$. For small values of A $M^U > M^B$, and for large values $M^U < M^B$. There exists a point at which M^U crosses M^B and this point is unique. If I add equations (1.16) and (1.17) I get equation (1.18), which implies that M^A, M^U and M^B intersect at one point, $\bar{A}_2$. If $A < \bar{A}_2$, M^U lies below M^A but above M^B, $M^A > M^U > M^B$, and conversely, if $A > \bar{A}_2$, $M^B > M^U > M^A$. □

Conditions (1.20)–(1.22) are depicted in Figure 1.3. Condition (1.20) describes the case when due to low social immigration costs the immigration quota is set high enough and it exceeds three other outcomes. In this case the quota will not be exhausted because less migrants wish to move under the free migration.

When condition (1.21) holds the social cost parameter is high enough and brings the immigration quota below the market level. This establishes the second key result of the paper. Given the finding of over-emigration, enforcement of the immigration quota by the host country benefits the welfare of the source country, because it reduces the volume of excessive migrants from $M^M - M^B$ to $M^A - M^B$, thus reducing the negative diaspora effect. Contrary to the well-acknowledged opinion that developed countries should accept more migrants to increase the welfare in the developing source region, if condition (1.21) holds, enforcement of the quota actually benefits the source country.

If condition (1.22) holds, the social immigration costs are too high and the quota is set too low. If the quota is enforced, there will be rationing of migrants and under-emigration with respect to what is best for the source country, the union and the market.

1.3.7 Comparison with brain drain

The brain drain literature finds that emigration of skilled workers decreases the welfare of those left behind and reduces economic growth in sending countries (Bhagwati and Hamada, 1974; Bhagwati and Hamada, 1982; Dustmann et al., 2011; Mountford and Rapoport, 2011). In comparison, this result establishes that if the welfare of the would-be migrants while they are in Country B is accounted for, the effect on stayers is zero. However, if the would-be migrants are excluded from the welfare, then for given M the marginal effect on stayers is negative.

The marginal effect of emigration on stayers and would-be migrants from equation (1.11) is:

$$L^B \frac{\partial w_L^B}{\partial m} + \left(H^B - m\right) \frac{\partial w_H^B}{\partial m} = 0. \tag{1.23}$$

The total effect is also zero:

$$\int_0^M L^B \frac{\partial w_L^B}{\partial m} + \left(H^B - m\right) \frac{\partial w_H^B}{\partial m}\, dm = 0. \tag{1.24}$$

Suppose now that a social planner knows M in advance and wishes to compute the welfare effect on stayers excluding the would-be migrants while they are still in Country B. The marginal effect is then:

$$L^B \frac{\partial w_L^B}{\partial m} + \left(H^B - M\right) \frac{\partial w_H^B}{\partial m} < 0 \quad \text{for} \;\; M > m. \tag{1.25}$$

For $M > m$ the marginal effect in equation (1.25) is negative, and for $M = m$ it is zero. The total effect is therefore negative and given by:

$$\int_0^M L^B \frac{\partial w_L^B}{\partial m} + \left(H^B - M\right) \frac{\partial w_H^B}{\partial m}\, dm < 0. \tag{1.26}$$

Unlike the brain drain literature, the non-positive effect on the welfare of stayers holds for the emigration of workers of any skill level and from any migrant sector of Country B: medical professionals, construction workers, computer scientists or cleaners.

1.4 Conclusions

In this paper I analyze the welfare effects of migration for three parties involved: migrants, receiving country and sending country. As stated in many studies, workers move in response to the wage differential between countries after deducing individual migration costs. Thus, when deciding to move, individual workers disregard their effect on the host and source countries' welfare as well as other migrants. The free migration level confronts the immigration quota of the receiving country; all prospective migrants move as long as the market level is below the quota, and the prospective migrants are rationed when the quota is below the market level.

In the absence of social immigration costs, immigration strictly benefits the receiving country. When these costs are not zero, the immigration quota is determined when the marginal increase in the host country workers' income equals the marginal increase in the social costs.

For the source country, emigration is found to decrease the output by the worker's wage. This, however, does not affect the income of stayers. I find that there is always over-emigration with respect to what is optimal for the source country because the individual decision rule does not account for the migrants' effect on their own income (diaspora effect), which is negative since migrants cluster in the same employment sector. Over-emigration is harmful to the source country welfare. Under certain conditions the negative impact of over-emigration can be reduced when the destination country enforces the quota.

The source country prefers to coordinate the immigration quota with the host country, because in the coordinated outcome of the political union the migrant workers stop being migrants and the negative diaspora effect is internalized. When the social immigration costs are zero, the union quota delivers the same outcome as the free market.

Bibliography

Akin, S., 2012. Immigration, fiscal policy, and welfare in an aging population. The B.E. Journal of Macroeconomics (advances) 12 (1).

Batista, C., Lacuesta, A., Vicente, P., 2012. Testing the 'brain gain' hypothesis: micro evidence from Cape Verde. Journal of Development Economics 97 (1), 32–45.

Beine, M., Docquier, F., Özden, Ç., 2011. Diasporas. Journal of Development Economics 95 (1), 30–41.

Benhabib, J., 1996. On the political economy of immigration. European Economic Review 40 (9), 1737–1743.

Bertoli, S., Brücker, H., 2011. Selective immigration policies, migrants' education and welfare at origin. Economics Letters 113 (1), 19–22.

Bhagwati, J., Hamada, K., 1974. The brain drain, international integration of markets for professionals and unemployment: a theoretical analysis. Journal of Development Economics 1 (1), 19–42.

Bhagwati, J., Hamada, K., 1982. Tax policy in the presence of emigration. Journal of Public Economics 18 (3), 291–317.

Bianchi, M., 2013. Immigration policy and self-selecting migrants. Journal of Public Economic Theory 15 (1), 1–23.

Billet, C., 2010. EC readmission agreements: A prime instrument of the external dimension of the EU's fight against irregular immigration. an assessment after ten years of practice. European Journal of Migration and Law 12 (1), 45–79.

Borjas, B., 1987a. Self-selection and the earnings of immigrants. American Economic Review 77 (4), 531–553.

Borjas, G., 1995a. The economic benefits from immigration. Journal of Economic Perspectives 9 (2), 3–22.

Borjas, G. J., 1987b. Immigrant, minorities, and labor market competition. Industrial and labor relations review 40 (3), 382–392.

Borjas, G. J., 1995b. The economic benefits from immigration. Journal of Economic Perspectives 9 (2), 3–22.

Borjas, G. J., 2011. Poverty and program participation among immigrant children. The Future of Children 21 (1), 247–266.

Bratsberg, B., Raaum, O., 2012. Immigration and wages: Evidence from construction. Economic Journal 122 (565), 1177–1205.

Burda, M., Wyplosz, C., 1992. Human capital, investment and migration in an integrated Europe. European Economic Review 36 (23), 677–684.

Card, D., 1990. The impact of the Mariel boatlift on the Miami labor market. Industrial and Labor Relations Review 43 (2), 245–257.

Carrington, W. J., Detragiache, E., Vishwanath, T., 1996. Migration with endogenous moving costs. American Economic Review LXXXVI, 909–930.

Cellini, R., 2007. Migration and welfare: a very simple model. Journal of International Development 19 (7), 885–894.

Clark, X., Hatton, T., Williamson, J., 2007. Explaining US immigration, 1971-1998. Review of Economics and Statistics 89 (2), 359–373.

Docquier, F., Rapoport, H., Salomone, S., 2012. Remittances, migrants' education and immigration policy: theory and evidence from bilateral data. Regional Science and Urban Economics 42 (5), 817–828.

Dustmann, C., Fadlon, I., Weiss, Y., 2011. Return migration, human capital accumulation and the brain drain. Journal of Development Economics 95 (1), 58–67.

Filer, R., 1992. The effect of immigrant arrivals on migratory patterns of native workers. In: Immigration and the Workforce: Economic Consequences for the United States and Source Areas. NBER Chapters. National Bureau of Economic Research, Inc, pp. 245–270.

Frontex, 2011. General report 2011. Tech. rep., Frontex.

Fuest, C., Thum, M., 2000. Welfare effects of immigration in a dual labor market. Regional Science and Urban Economics 30 (5), 551–563.

Gera, S., Songsakul, T., 2007. Benchmarking Canada's performance in the global competition for mobile talent. Canadian Public Policy – Analyse de politiques 33 (1), 63–84.

Giordani, P. E., Ruta, M., 2011. The immigration policy puzzle. Review of International Economics 19 (5), 922–935.

Heckman, J. J., 1979. Sample selection bias as a specification error. Econometrica 47 (1), 153–161.

Ketkar, S. L., Ratha, D., 2010. Diaspora bonds: tapping the diaspora during difficult times. Journal of International Commerce, Economics and Policy 1 (2), 251–263.

LaLonde, R. J., Topel, R. H., 1991. Immigrants in the American labor market: quality, assimilation, and distributional effects. American Economic Review 81 (2), 297–302.

Marchiori, L., Shen, I.-L., Docquier, F., 2013. Brain drain in globalization: a general equilibrium analysis from the sending countries' perspective. Economic Inquiry 51 (2), 1582–1602.

Mattoo, A., Neagu, I. C., Özden, Ç., 2008. Brain waste? Educated immigrants in the US labor market. Journal of Development Economics 87, 255–269.

Mavroudi, E., Warren, A., 2013. Highly skilled migration and the negotiation of immigration policy: Non-EEA postgraduate students and academic staff at English universities. GEOFORUM 44, 261–270.

Miller, P. W., 1999. Immigration policy and immigrant quality: the Australian points system. American Economic Review 89 (2), 192–197.

Mincer, J., 1978. Family migration decisions. Journal of Political Economy 86 (5), 749–773.

Mountford, A., 1997. Can a brain drain be good for growth in the source economy? Journal of Development Economics 53 (2), 287–303.

Mountford, A., Rapoport, H., 2011. The brain drain and the world distribution of income. Journal of Development Economics 95 (1), 4–17.

OECD, 2011. International Migration Outlook. OECD Publishing, Paris.

OECD, 2012. International Migration Outlook. OECD Publishing, Paris.

OECD, 2013. Recruiting Immigrant Workers: Germany 2013. OECD Publishing.

Ortega, F., Peri, G., 2009. The causes and effects of international migrations: evidence from OECD countries 1980-2005. NBER Working paper No. 14833.

Ostrovsky, Y., 2012. The dynamics of immigrant participation in entitlement programs: evidence from Canada, 1993–2007. Canadian Journal of Economics/Revue canadienne d'économique 45 (1), 107–136.

Rabe, B., 2011. Dual-earner migration. earnings gains, employment and self-selection. Journal of Population Economics 24, 477–497.

Razin, A., Sadka, E., 1999. Migration and pension with international capital mobility. Journal of Public Economics 74 (1), 141–150.

Schiff, M., 2002. Love thy neighbor: trade, migration, and social capital. European Journal of Political Economy 18 (1), 87–107.

Stark, O., Casarico, A., Devillanova, C., Übelmesser, S., 2012. On the formation of international migration policies when no country has an exclusive policy-setting say. Regional Science and Urban Economics 42 (3), 420–429.

Stark, O., Helmenstein, C., Prskawetz, A., 2004. Rethinking the brain drain. World Development 32 (1), 15–22.

Storesletten, K., 2000. Sustaining fiscal policy through immigration. Journal of Political Economy 108 (2), 300–323.

UN, 2011. International migrant stock by origin: the 2010 revision. Population Division, Migration Section.

Zimmermann, K. F., 1995. Tackling the European migration problem. Journal of Economic Perspectives 9 (2), 45–62.

Appendix A

A.1 Key migration statistics

Table A.1.1: *Standardized migration statistics.*

	AUT	BEL	CZE	HUN	GBR	FIN	FRA	USA	CHE	DEU	GRC	LUX	DNK	IRL	SWE	ITA	NOR	NLD	PRT	ESP
Permanent inflows and outflows in 2010, per one thousand of local population*																				
Inflows	11.7	10.4	2.9	2.4	8.1	3.4	2.2	3.4	17.2	8.4	2	31.5	6	3.9	8.4	7.1	13.3	6.6	2.8	9.4
Outflows	7.9	.	1.4	0.6	3.3	0.6	.	.	8.4	8.2	4.2	15.2	4.9	8.4	2.4	0.5	4.6	2.4	.	7.3
Inflows in 2011 by category of entry, %																				
Work	1.4	18.3	.	.	33.1	5.8	11.9	6.4	2.1	9.0	.	.	19.6	16.3	5.7	40.5	5.1	10.9	21.9	29.9
Free movements	63.7	39.6	.	.	17.4	39.0	30.3	0.0	71.4	59.9	.	.	50.9	71.8	35.9	28.2	67.4	56.9	36.3	49.9
Accompanying family	0.9	0.0	.	.	14.6	0.0	0.0	7.8	0.0	0.0	.	.	5.9	4.0	0.0	1.2	0.0	0.0	0.0	0.0
Family	23.2	36.2	.	.	11.8	34.3	42.9	66.3	18.8	24.7	.	.	12.3	7.0	39.6	27.4	18.0	21.7	35.3	18.7
Humanitarian	10.3	5.9	.	.	1.2	17.4	5.4	13.1	5.8	5.3	.	.	5.1	0.9	18.7	1.3	9.5	10.5	0.1	0.2
Other	0.5	0.0	.	.	21.9	3.6	9.6	6.4	2.0	1.1	.	.	6.2	0.0	0.0	1.5	0.0	0.0	6.3	1.2
Stock of migrants in 2010, % of population																				
Foreign-born	15.7	13.9	6.3	4.5	11.5	4.6	8.6	12.2	26.6	13	10.9	37.6	7.7	17.3	14.8	8	11.6	11.2	.	14.5
Foreign nationals	11.1	9.8	4	2.1	7.4	3.1	6	7	22.1	8.3	7.1	44.1	6.2	.	6.8	7.6	7.6	4.6	4.2	12.4
Unemployment rate in 2011, %																				
Native-born	3.4	5.8	6.8	11.1	8	7.6	8.5	9.2	3.1	5.4	17.4	3.4	6.9	14.1	6	8	2.7	3.8	13	19.5
Foreign-born	8.2	15.1	8	9.5	9.4	15.2	15.1	9.1	6.8	9.5	22.2	6.3	14.5	17.3	16	11.7	7.7	9.2	16.9	31.5
Share of foreign-born in tot. empl., %, 2010	16.5	13.0	2.8	2.1	13.3	3.8	11.1	15.8	27.2	14.3	11.5	48.7	9.4	17.2	14.4	11.8	9.5	11.3	9.1	17.0
Difference in education level shares: recent migrants (2000–2010) minus young resident workers**																				
Low	16.4	18.0	6.9	2.3	14.1	29.4	23.9	26.4	11.8	17.7	41.9	1.0	.	9.0	20.7	28.9	.	17.6	11.1	13.3
Medium	-21.2	-7.3	-4.2	-8.6	6.1	-6.6	-12.4	-12.0	-22.0	-31.2	-8.8	-26.5	.	3.8	-28.8	-12.0	.	-10.1	8.9	13.3
High	4.7	-10.7	-2.8	6.4	-20.2	-22.8	-11.6	-14.4	10.1	13.5	-33.1	25.4	-11.8	-12.8	8.1	-17.0	-11.5	-7.5	-20.0	-26.6

Notes: * National definitions of migrant. ** “Low” refers to less than upper secondary attainment, “Medium” to upper secondary and post-secondary non-tertiary, “High” to tertiary. A dot means that the data are unavailable or the estimate is unreliable.
Source: author’s illustration using OECD (2011) and OECD (2012) data.

Table A.1.2: *Migrants' employment by sector and occupation.*

	Austria	Belgium	Czech Rep.	Hungary	UK	Finland	France	US	Switzerland	Germany	Greece	Luxembourg	Denmark	Ireland	Sweden	Italy	Norway	Netherlands	Portugal	Spain
	Distribution of migrants across employment sectors in 2011*, %																			
Agriculture and fishing	0.9	0.6	1.2	.	0.5	2.6	1.3	2.2	1.2	0.7	8.9	.	2.4	2.2	0.6	4.0	.	1.7	.	5.7
Mining, manufacturing and energy	17.8	13.3	31.1	23.9	10.6	15.3	11.7	12.9	18.4	25.6	13.4	6.8	13.9	17.1	13.2	20.6	12.0	14.5	14.1	9.3
Construction	11.8	9.6	10.1	7.9	5.7	6.1	12.3	9.3	8.0	7.0	19.2	9.9	1.8	4.2	4.1	13.6	6.2	4.2	9.9	10.5
Wholesale and retail trade	14.6	11.8	15.1	16.1	13.2	12.8	11.7	13.9	14.2	12.7	14.6	10.8	11.8	16.3	11.1	10.1	11.2	12.6	13.3	13.6
Hotels and restaurants	12.1	8.2	5.6	5.7	9.2	8.5	7.0	10.5	7.7	8.8	12.1	5.9	7.5	11.7	7.2	8.8	6.8	6.8	10.5	16.1
Education	4.3	5.5	5.2	9.8	8.2	8.1	5.2	5.8	5.3	4.5	1.5	4.1	9.5	4.9	11.3	1.9	6.5	6.6	9.1	2.2
Health	9.5	10.9	4.6	9.4	15.6	11.7	11.6	12.1	13.1	10.7	3.2	7.4	21.5	14.4	19.3	4.8	24.3	16.4	7.2	5.0
Households	.	2.2	.	.	0.4	.	5.4	1.4	1.3	1.1	14.7	3.5	.	1.0	.	17.0	.	.	5.1	14.9
Administrative	2.6	9.9	3.1	4.5	4.1	2.5	6.4	2.4	2.2	2.4	0.9	14.4	3.2	2.2	3.7	1.4	3.0	6.3	7.1	2.0
Other services	25.8	28.0	23.9	20.1	32.4	32.3	27.5	29.5	28.6	26.6	11.5	37.0	28.0	26.0	29.4	17.7	28.5	30.8	22.2	20.6
	Difference in occupation shares: recent migrants (2000–2010) minus young resident workers																			
Elementary occupations	17.1	11.4	9.4	1.8	10.6	17.9	15.2	14.6	4.7	14.3	38.7	3.1	19.2	17.7	12.5	34.2	14.2	16.2	24.6	34.0
Clerks, service workers, skilled trades, machinery operators	-1.2	-1.7	5.4	-4.2	3.3	0.6	3.0	1.7	-5.3	2.6	0.8	-7.4	3.8	18.5	-1.4	-1.0	5.8	7.4	5.5	2.9
Professionals, senior officials and managers	1.9	-5.6	-0.8	11.7	-7.5	-6.0	-2.9	-16.3	11.9	-2.4	-25.4	14.3	-8.6	-28.8	-1.1	-9.3	-3.8	-13.8	-18.3	-20.3
Technicians and associate professionals	-17.8	-4.1	-14.1	-9.3	-6.5	-12.6	-15.3	0.0	-11.3	-14.5	-14.1	-10.0	-14.4	-7.4	-9.9	-23.9	-16.3	-9.7	-11.8	-16.5

Notes: * ISIC classification, rev.3. A dot means that the data are unavailable or the estimate is unreliable.
Source: author's illustration using OECD (2012) data.

Chapter 2

Immigration Policy Index

Abstract

I construct an immigration policy index which is heterogeneous across destination-origin country pairs and variant over time. This index is based on three types of entry visa restrictions: visa required, visa not required for short stays and visa not required at all. When estimated in levels, visa exempt country pairs account for around 15% more migrants than their counterfactual. I show that the effects of migration determinants vary by the type of visa restrictions. Further, I identify country pairs which changed their visa regime during 2000–2010 and find that the weakening of visa policy is associated with a 10% increase in migrant stocks and a significant shift toward male and less skilled migration from policy affected source countries. In contrast, the tightening of visa policy is not related to a significant change in migrant stocks, their gender or skill composition.

Keywords: immigration policy; visa; difference-in-difference estimation; policy quasi-experiment; group heterogeneity; diaspora effect

JEL Classification: F22; K37; F66; R23

This paper is published as CERGE-EI working paper No. 523. The earlier version was circulated as MPC Analytical and Synthetic Note 2013/02. I am grateful to Byeongju Jeong, Jan Kmenta, Nikolas Mittag, Randy Filer, Alexei Savvateev and Sherzod Tashpulatov for helpful comments. I am thankful to Ladislav Bednář from ČSA for kindly providing access to data. I benefited from the discussions with summer school lecturers at EUI in Florence (2012), University of Rethymnon (2013), University of Milan (2013), Parthenope University in Naples (2014) and MPI in Jena (2014).

2.1 Introduction

Immigration policy is one of the key determinants of international migration. It consists of rules which govern the admission of foreigners, their access to the labor market, healthcare, welfare, voting, and family reunion. This complex set of rules also means that it is difficult to measure immigration policy empirically and calls for more research in the area.

In this paper I suggest using entry visa rules to measure the tightness of immigration policy. Entry restrictions affect the amount of migrants through information and feasibility channels. In the absence of any visa barriers between two countries, workers can move freely, disseminating information about employment opportunities abroad (information channel). In response to a wage gap, workers take up jobs abroad because they are not restricted in mobility by visas (feasibility channel).

In the presence of visa restrictions, workers' mobility and access to information is restricted because they need to apply for an entry clearance (visa) prior to traveling. In many cases a positive outcome of an application is not guaranteed because job-seeking alone is not a sufficient reason to obtain a visa. Even though information on better paying jobs abroad might be generally available, workers cannot take them due to travel and employment restrictions. All else being equal, firms might be reluctant to hire migrant workers because of extra paperwork.

The immigration policy index is constructed for all countries and territories in the world as of March 1998 and November 2009. This index is heterogeneous across destination and origin countries as well as over time. I find that country pairs with simplified entry restrictions (visa partially required or visa not required) account for about 15% more migrants than pairs with visa required status. Also, these migrants are more likely to be males and less educated.

I further set up a quasi-policy experiment which tracks down policy shifters for which immigration policy changed between March 1998 and November 2009. I find that after visas were abolished, the stocks of migrants in affected country pairs increased 10% relative to their counterfactual. The increase was predominantly in male and less educated migration. In contrast, the introduction of visas was not associated with a statistically significant change in the stock of migrants, their gender or education composition for the period considered.

Although most studies surveyed do somehow acknowledge the presence of immigration policy, more research needs to address its complex nature. For the succinctness of exposition I summarize key studies in the literature in Table B.2.1. I only review papers that analyze country level aggregate data. In such studies a unit of observation is a destination-origin country pair (dyad) in year t.

The determinants of international migration can be broken down into four major groups: economic incentives, demographic factors, linguistic and cultural proximity, and institutional factors. While the first three groups have been extensively studied in the literature, immigration policy remains under-researched mainly due to the lack of comparable cross-country data. The role of these factors is also studied in wider literature on the determinants of international trade (Head and Mayer, 2015) and economic growth (Alesina et al., 2003).

The idea of using visa restrictions to quantify institutional barriers to mobility is not new. Hobolth (2014) constructs a European visa database for whether a sending country needs a visa to the EU destinations. Glaesser and Kesler (2013) consider visa restriction as an obstacle to tourism and construct an aggregate index of visa openness for each country. Based on these data, Neumayer (2010) estimates that the presence of visas is associated with a 52–63% reduction in tourism related travel. Using data from expert surveys, Huddleston et al. (2011) creates 7 aggregate indexes that compare the national policies of 31 developed destinations on family reunion, access to labor market, education, nationality, and voting.

Since most authors analyze immigration into the OECD countries, they include Schengen or EU dummies. Grogger and Hanson (2011) and González and Miles-Touya (2014) also add visa waiver dummies. To control for the tightness and skill selectivity of immigration policies, Beine et al. (2011), Grogger and Hanson (2011), and Docquier et al. (2012) use the shares of refugees and asylees in stocks. Mayda (2010) and Ortega and Peri (2013) follow a different approach. Based on destination country legislation, they construct aggregate country-specific indexes to proxy for the tightness of immigration policies. Palmer and Pytlikova (2013) and Kahanec et al. (2014) develop an index of labor market access laws for migrants within the EU.

I contribute to the literature in two ways. First, I expand the classification of visa waivers to three categories: visa required, visa partially required and visa not required. A dummy variable from existing studies becomes a categorical variable in this study, thus generating more variation. Second, using IATA (1998) and IATA (2009) data, I create this categorical variable for all countries and territories in the world at two points in time: March 1998 and November 2009. This enables me to identify policy shifters and track changes in the stocks of migrants, their gender and skill composition relative to non-shifters.

The emphasis on world migration extends many existing studies which mainly emphasize migration into the developed OECD countries. Belot and Ederveen (2012) analyze migration only within the developed world, claiming that the mechanism behind North↔North migration is somehow different. However, the UN (2011) estimate that North↔North migration accounts for 25% of world migration, whereas migration within the developing countries of the South alone amounts to 35% of the world total. By analysing world migration data, I learn about the data generating mechanism behind migration for all countries, not only the developed ones.

The paper is structured as follows. I first construct a simple theoretical model of the determinants of migration. Sections 2.3 and 2.4 describe the construction of the index and policy quasi-experiment. Next, I describe the data used, set up and estimate the empirical model, check its robustness and discuss the estimation results.

2.2 The model of the determinants of migration

There is an active debate in the literature as to whether the utility function of a representative worker is linear or log-linear in wage gain from emigration. Suppose a worker's wage at origin is 100 units and his wage at destination is 120 units. For a linear utility function, the net gain from emigration (assuming zero migration costs) is 20 units. For a log-linear utility, this increase is 20% of the current wage. A worker with a linear utility function cares about an absolute wage gain, whereas a worker with a log-linear function cares about the magnitude of wage increase relative to the current level. Grogger and

Hanson (2011) elaborate on differences between these two functional forms.

In line with the debate, I develop two versions of the model which differ in their approach to individual migration costs. In model one the costs are expressed in terms of units and in version two they are modeled in terms of time.

2.2.1 Version one: additive migration costs

A developed country A with population N^A receives migrants from a developing country B with population N^B. In each country the population consists of skilled and unskilled workers, denoted by H and L, with respective shares in population α^A and α^B. Wages are assumed to be exogenous and satisfy the following inequality:

$$w_H^A > w_H^B > w_L^A > w_L^B. \tag{2.1}$$

Worker k emigrates if inequality (2.2) holds, and stays otherwise:

$$w_s^A - w_s^B > C_{ks}, \tag{2.2}$$

where C_{ks} stands for broadly defined migration costs of an individual k of skill type s, $s = L, H$. The costs are assumed to have the following additive structure:

$$C_{ks} = D - S - L - H - I + \nu_{ks}, \tag{2.3}$$

where D is distance between countries; S is the stock of migrants from country B in country A; L measures language similarity between A and B; H measures historic and cultural proximity between A and B; I measures the tightness of immigration policy of country A with respect to B; ν_{is} is a random variable. Since there are only two countries, I omit country superscripts.

Under the assumption of uniform distribution, $\nu_{is} \sim U[0,1]$, the stock of migrants of skill s, M_s, expressed as the share of N^B is:

$$\begin{aligned} \frac{M_s}{N^B} &= Prob\left(w_s^A - w_s^B - D + S + L + H + I > \nu_{ks}\right) = \\ &= w_s^A - w_s^B - D + S + L + H + I. \end{aligned} \tag{2.4}$$

Since equation (2.4) holds for both skill types, the total stock is:

$$\frac{M_L + M_H}{N^B} = w_L^A - w_L^B + w_H^A - w_H^B - 2D + 2S + 2L + 2H + 2I. \tag{2.5}$$

Wages w_s^A and w_s^B are not observed in data, but they can be inferred from average wages, the Gini index and average years of schooling through equations (B.18) and (B.19) derived in Appendix B.1.

2.2.2 Version two: multiplicative migration costs

The setup is the same as in version one, except individual migration costs C_{ks} are in multiplicative form. The decision rule (2.2) thus becomes:

$$\frac{w_s^A}{C_{ks}} > w_s^B, \tag{2.6}$$

where C_{ks} is a non-negative random variable which plays the role of an individual discount factor. For workers with low C_{ks} the effective wage abroad, $\frac{w_s^A}{C_{is}}$, exceeds the effective wage of workers with high C_{ks}. All else being equal, the effective wage favors individuals who speak foreign languages, have relatives living abroad and quickly adjust to new living conditions.

Individual costs are assumed to have multiplicative form:

$$C_{ks} = \frac{D}{S \cdot L \cdot H \cdot I} \cdot \nu_{ks}, \tag{2.7}$$

where D, S, L and I are positive continuous variables defined in version one. ν_{ks} is a random variable from the exponential family of distribution functions, $F(\nu) = \alpha(\nu)^\rho$. Parameters α and ρ are jointly determined so that $F(\cdot)$ satisfies the definition of a distribution function.

The stock of migrant workers of skill type s is:

$$\frac{M_s}{N^B} = Prob\left(\frac{w_s^A \cdot S \cdot L \cdot H \cdot I}{w_s^B \cdot D} > \nu_{ks}\right) = \alpha\left(\frac{w_s^A \cdot S \cdot L \cdot H \cdot I}{w_s^B \cdot D}\right)^\rho. \tag{2.8}$$

Taking the logarithm of both sides I obtain:

$$\ln \frac{M_s}{N^B} = \ln \alpha + \rho \left(\ln w_s^A + \ln S + \ln L + \ln H + \ln I - \ln w_s^B - \ln D \right). \tag{2.9}$$

In equation (2.9) all variables are in logarithms, whereas in equation (2.5) the variables are in levels. In order to choose between these two competing equations, I will apply the PE test (Kmenta, 1990, pp. 521–522) in Section 2.6.

2.3 Policy index design

The purpose of constructing the immigration policy index is to rank in a consistent manner the tightness of entry rules for all countries and territories in the world. The simplification of entry rules reduces institutional barriers, facilitates a better job search and lowers mobility costs, allowing people to respond to economic incentives more elastically. It is interesting to estimate the extent to which these entry restrictions affect migrant stocks, their gender and skill composition. Empirical studies that do not account for immigration policy might produce unreliable estimates due to omitted variable bias.

Imagine that by default every sovereign country demands a visa from arriving foreign nationals. At the same time every destination has a number of source countries with which it has friendly relations and thus simplified visa regimes. Governed by data, I distinguish three major categories of entry rules: "visa is required," "visa is not required for a stay shorter than n days," and "visa is not required". Examples of country pairs in each category are given in Table B.2.3.[1]

The default state "visa is required" is when a person prior to commencing their journey has to contact the nearest embassy of the destination country (or other country liable to issues visas on its behalf), submit a visa application in person, online or by ordinary mail. Usually a letter of invitation or sponsorship from a hosting institution, company or family at destination is required. This process is time consuming and in many instances it is advised to start the application process at least a month prior to the planned travel

[1] A traveler is assumed to hold a normal passport (not consular, diplomatic, service, business or special passport), travel alone as a tourist for a very short stay from a country of origin and hold no valid visas to other destinations.

date. Afghanistan, Myanmar, and Turkmenistan are examples of countries that demand visas from every nationality.

The category "visa is not required for stays shorter than n days," also referred to in this text as "visa is partially required," is when a host country allows certain nationals to enter without a visa for stays not exceeding a certain number of days. For 2010 this limit ranges from 7 days for Togo←New Zealand (tourists in Togo from New Zealand) to 365 days for Palau←Micronesia. The two most frequently observed durations are 90 and 30 days. Often the duration of an allowed visa free stay varies according to the purpose of travel: tourist- or family-related stays are on average allowed for longer than business-related stays. Often a traveler is required to hold a return ticket, sufficient funds for the duration of stay and produce evidence of a reservation of accommodation. This category also combines countries that issue visas upon arrival for a fee.

The US visa waiver and Australia ETA and eVisitor online visas application programs fall into this category. These programs allow certain nationals to apply for travel authorisation online and avoid lengthy application procedures and enhanced security checks.

I group allowed durations of stay and provide examples of country pairs in each group in Table B.2.4. I also regress $\ln(stock_{ijt})$ on a set of duration group dummy variables to learn if a longer duration of stay can be associated with larger migrant stock. Indeed, this is the case as the estimates in Table 1 suggest.

Table 1: Differences in the means of $\ln(stock_{ijt})$ *by the duration of visa-free stay, pooled sample.*

Variable	Estimate	S.E.
$[3, 30)$ days	*base category*	
30 and 31 days	0.332	(0.58)
$[45, 90]$ days	1.106**	(0.51)
$[120, 365]$ days	1.968***	(0.76)
y10	-0.041	(0.14)
cons	4.573***	(0.49)

Notes: The number of observations is 8175, adj. $R^2 = 0.024$. * p<0.1, ** p<0.05, *** p<0.01.

The least restrictive entry category is "visa is not required," when no limit is imposed on the duration or purpose of stay. Quite often visa-free travel is associated with simplified access to the local labor market. For 2010 this regime is established between the USA↔Canada, Australia↔New Zealand, EEA countries, most republics of the former Soviet Union, Gulf Cooperation Council countries, Algeria↔Morocco↔Tunisia and Uganda↔Eritrea↔Kenya, to mention some of the most prominent examples. The existence of a visa-free regime is associated with regional integration, enhanced bilateral trade and development programs.

A note should be made on the classification of countries' overseas territories.[2] From the perspective of a destination country, mainland and overseas territories are considered separate items because in most cases such territories share different immigration policies than their mainland countries. For example, Puerto Rico and the US Virgin Islands' visa waiver programs are less restrictive than the US mainland program. Many of the UK overseas territories are popular tourist destinations and have more welcoming immigration policies than the UK mainland. For example, the Turks and Caicos Islands and Bermuda issue visas upon arrival for up to one month and six months, respectively, to most nationalities. French Polynesia allows only a 90-day visa waiver for the EU countries, compared to visa-free entrance granted by mainland France.

From the perspective of a sending country, overseas territories and the mainland share a similar, if not the same, passport and are thus considered one sending country. For example, for most destinations in the world a British passport, which is shared by the citizens of the UK and the British overseas territories, grants equal immigration rights irrespective of the endorsement in the passport.[3] Also, for most destinations equal rights hold for passports issued by the source country mainland and overseas territories.

An apparent advantage of the constructed index is that it provides variation across country pairs, over time and can be constructed for all country pairs. This is a relative improvement over the indexes of Ortega and Peri (2013), Mayda (2010) and Palmer and Pytlikova (2013). The visa index also has a clear intuitive design and straightforward interpretation in regression analysis. This extends the analyses of Docquier et al. (2012) and Grogger and Hanson (2011) in a way that the shares of females or refugees in stocks

[2]For the correctness of terminology, each country has its own term for overseas territories: unincorporated territory (US), constituent country (the Netherlands), overseas department / collectivity/ sovereignty (France) and autonomous country (Denmark).

[3]Possible endorsements are: British Citizen, British Overseas Territories Citizen, British national (overseas), British Overseas Citizen, British Protected Person and British Subject.

are the outcome of demand and supply equations. Individuals decide to apply for entry clearance and then a destination country decides whether to grant an entry permit.

A disadvantage of the created index is the ignorance of migrants' rights to employment, access to various benefits, healthcare, which are not captured by entry visa rules. However, it normally holds that less restrictive entry rules are associated with more rights granted to migrants.

2.4 Policy quasi-experiment setup

I compile the immigration policy index for March 1998 and November 2009. The stock data are observed for the middle of 2000 and 2010. The purpose of this policy quasi-experiment is to identify country pairs for which the policy index was changed during 2000–2010 and investigate how these changes reflected on the stock of migrants and their composition relative to country pairs with an unchanged policy index.

There are two types of policy changes: policy weakening (up-shifters) and policy tightening (down-shifters). The treatment group consists of shifter country pairs and the control group is composed of non-shifter pairs. The treatment effects are heterogeneous, because there are three types of up-shifter country pairs and the same number of down-shifter pairs. These types are explained further below.

There are two underlying hypotheses to be tested. Hypothesis one states that the introduction of entry visas decreases the amount of migrants from affected source countries. According to hypothesis two, the abolition of visas increases the stocks of migrants from the countries in question. The span of 10 years is assumed to provide sufficient time for policy change to take effect. Finally, it is worth investigating how the visa rules affect the gender and skill composition of migrant stocks. Previous studies have documented that stricter immigration policies are associated with more skilled migrants (Grogger and Hanson, 2011; Beine et al., 2011).

Table 2 shows that for about 19% of country pairs the visa regime was changed between 2000 and 2010. Visa policy was weakened for 13.5% and tightened for 5.5% of country pairs. The weakening of policy means one of the following two statements:

1. The visa required regime was changed to visa partially required or visa not required;

2. Visa partially required was changed to visa not required.

Symmetrically, the tightening of visa policy implies one of the following two statements:

1. Visa partially required was changed to visa required;

2. Visa not required was changed to visa partially required or visa required.

Table 2: Tabulation of policy changes in 2000–2010.

		2010 Visa required	 Visa partially required	 Visa not required	 Total
2000	Visa required	24722	4741	200	29663
	Visa partially required	1914	8623	712	11249
	Visa not required	113	256	549	918
	Total	26749	13620	1461	41830

The data in Table 2 show that all types of policy changes took place during 2000–2010. The most frequent policy change is the move from visa required to visa partially required (4741 country pairs). This includes the extension of the US visa waiver program, ETA and eVisitor programs in Australia and Federal Skilled Worker Program in Canada. The EU granted partial visa waivers to Bolivia, Costa Rica, and most of the British overseas territories.

For most countries there is a clear tendency to become more open to immigration. Since I have a 217-by-202 matrix of country pairs over two points in time, it is difficult to summarize the dynamics of immigration policy for each country pair without aggregation. For each destination country i and $t = \{2000, 2010\}$ I compute $\alpha(i, t, vp)$, the share of origin countries in each visa category vp, $vp = \{0, 1, 2\}$. Obviously, $\sum_{vp=0}^{2} \alpha(i, t, vp) = 1 \; \forall i$ and t. For given i, $d(i, vp)$ denotes the time change of each share in group vp, $d(i, vp) = \alpha(i, t = 2010, vp) - \alpha(i, t = 2000, vp)$, leading to the following identity:

$$d(i, vp = 0) + d(i, vp = 2) = -d(i, vp = 1). \tag{2.10}$$

Since the differenced shares are linearly dependent (equation 2.10), it is sufficient to consider only two arbitrary shares. In Figure 1, I plot $-d(i, vp = 0)$ against $d(i, vp = 2)$.

In this figure, moving right along the horizontal axis means that a destination country weakened its visa policy by expanding the visa not required category and narrowing the visa required and/or visa partially required categories. Moving up the vertical axis means that a destination country weakened its visa policy by shrinking the visa required category and expanding the visa partially required and/or visa not required categories.

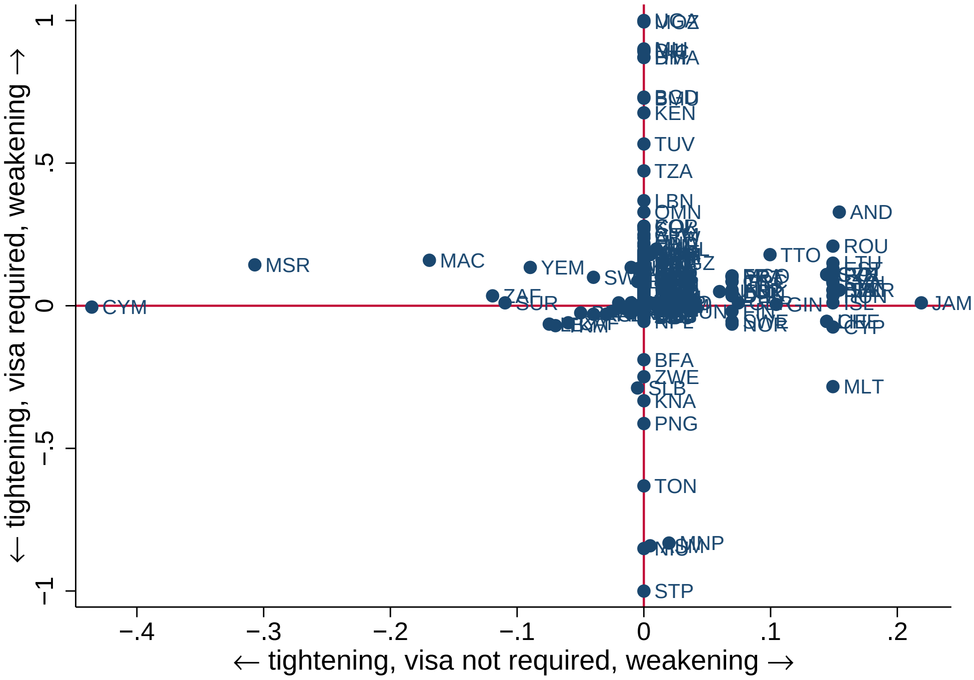


Figure 1: Scatter plot of policy changes in 2000–2010.

All visa policy changes can be structured as follows. Along the horizontal axis in Figure 1 are located countries that expanded or contracted the visa not required category by reshuffling between the visa required or visa partially required categories. Jamaica, Romania and Andorra expanded the visa not required and narrowed the visa required categories.

Along the vertical axis are the countries that altered the visa required category by expanding or narrowing the visa not required or visa partially required categories. Oman, Lebanon, and Kenya became more liberal by narrowing visa required and expanding visa partially required categories. The cluster of countries in the middle did not change their entry policies.

2.5 Data description

2.5.1 Construction of variables

This section rationalizes data sources used and explains the construction of variables. Table B.2.2 summarizes the definitions of variables and data sources.

Visa policy. Raw data for the visa policy dummies come from IATA (1998) and IATA (2009). Each of these two sources is a paper-back manual with various information for each country and territory in the world. For 1998 and 2009 I create a matrix of destination-origin country pairs.[4] Due to missing information on Georgia, Armenia, Moldova, Tajikistan and Congo (Brazzaville) in 1998, these entries are excluded from estimation in both years.

Migrant stocks. The decision to use stock but not flow data is motivated by wider country coverage. Flows and stocks are linked through the law of motion. Assuming zero stocks at $t = 0$, stocks at $t + 1$ are total flows until t adjusted for the rates of out-migration, naturalization and death. Changes in entry visas should be reflected in flow data immediately and in stock data with a sort delay.

There are several sources of comprehensive macro level migration data. OECD (2013), UN (2013) and Özden et al. (2011) provide stocks disaggregated by gender. Of these three, OECD (2013) is the most frequently cited source; however, it contains information only on the OECD destinations. Artuc et al. (2013), Defoort (2008) and Brücker et al. (2013) supply stock data by educational attainments for selected destinations. I use UN (2013) and Brücker et al. (2013) data because they provide the most up-to-date and extensive geographical coverage at the time of writing this paper.

Wage data. I impute unskilled and skilled wages from GDP per capita (Feenstra et al., 2013), the Gini index (Solt, 2014) and years of schooling data (Barro and Lee, 2013) using equations (B.18) and (B.19) in Appendix B.1. Grogger and Hanson (2011) mention a similar method as one possible imputation technique.

Cultural proximity. I use all six dummy variables from Head et al. (2010) data to proxy for cultural and historic proximity of a country pair. Two apparent issues arise

[4]The created matrix is not symmetric due to the different treatment of dependent territories as receiving and sending countries.

here. First, some of them are highly correlated, potentially leading to multicollinearity. Second, I am interested in the overall degree of cultural and historic similarity, irrespective of whether this comes from a colonial past, having a common coloniser or being in the same country. To address both issues, I perform principal component analysis, which reduces the dimensionality to three components:[5]

$$
\begin{aligned}
pc_1 &= 0.06 \cdot D_{contig} + 0.64 \cdot D_{colony} - 0.05 \cdot D_{comcol} + 0.39 \cdot D_{curcol} + 0.66 \cdot D_{col45} + 0.04 \cdot D_{smctry}, \\
pc_2 &= 0.57 \cdot D_{contig} + 0 \cdot D_{colony} + 0.49 \cdot D_{comcol} - 0.03 \cdot D_{curcol} - 0.04 \cdot D_{col45} + 0.65 \cdot D_{smctry}, \\
pc_3 &= -0.54 \cdot D_{contig} + 0 \cdot D_{colony} + 0.72 \cdot D_{comcol} + 0.43 \cdot D_{curcol} - 0.03 \cdot D_{col45} - 0.05 \cdot D_{smctry},
\end{aligned}
$$

where D_{contig} is a dummy variable for sharing a common border, D_{colony} stands for the same colony, D_{comcol} denotes a common colonizer, D_{curcol} is a dummy for currently being in a colony, D_{col45} is a binary variable for a colony after 1945 and D_{smctry} is a dummy variable for the same country.

These components explain more than 70% of variation in the original six variables. Component pc_1 assigns large weight to being in a colony. Component pc_2 emphasizes sharing the same border, being in the same country or sharing a common colonizer. Finally, component pc_3 favors contiguity, common colonizer, and current colony.

Language similarity. Using Ethnologue database (Lewis et al., 2013) I tabulate the 3 most frequently spoken languages for each country.[6] In some cases this includes official, regional and minority language (including those spoken by migrants). Then, I created a dummy variable that equals 1 if a country pair shares at lest one common spoken language. Out of 217×202 country dyads, 17.27% share a common language.

This approach slightly differs from existing studies. Head et al. (2010) create a widely circulated dataset with a dummy variable for a common official language. In many cases,

[5]Similar types of aggregation are frequent in cross-country studies: Melitz and Toubal (2014) create an aggregate index of a common language, Alesina et al. (2003) build measures of within country fractionalization based on the degree of linguistic, ethnic, and religious diversity of a country. In studies based on microdata, Vyas and Kumaranayake (2006) construct an aggregate index of individual socio-economic status, and Greene (2012) (example 4.12 on p. 93) constructs an index of online movie popularity.

[6]Ethnologue database alongside with CIA factbook, Encyclopedia Britannica and Wikipedia are also cited by Melitz and Toubal (2014), Belot and Ederveen (2012), Head and Mayer, 2015 and Alesina et al. (2003).

this does not capture regional language variation within a country.[7] To address the issue, Melitz and Toubal (2014) create three more variables: common spoken language and common native language (measured as the shares of speakers in a country pair, based on microdata) and linguistic proximity (based on assignment to branches in a language tree or the degree of similarity of a set of words).

2.5.2 Analysis of means of stocks

The collected dataset is a short panel of 217 destination and 202 origin countries and territories over two years (2000 and 2010). The unit of observation is a destination-origin country pair at time t, ijt, where i denotes a destination country and j stands for a source country. There are seven sources of variation in the data: across destination countries (42.7% of total variation), across source countries (17% of total variation), over time (0.3% of total variation) and combinations of the three (40% of total variation). The share of missing observations in the stock variable is 70% and the share of zeros is 2%.

I begin by describing the key variable $stock_{ijt}$ over two years and across three groups: visa required ($vp = 0$), visa not required for n days ($vp = 1$), and visa not required ($vp = 2$). Table 3 provides basic descriptive statistics on the pooled sample by visa group. The mean levels of groups are $\bar{x}(vp = 0) = 14437$, $\bar{x}(vp = 1) = 8932$ and $\bar{x}(vp = 2) = 43265$. The respective median values are $\tilde{x}(vp = 0) = 80$, $\tilde{x}(vp = 0) = 170$, and $\tilde{x}(vp = 0) = 1679$. Since the ordering of means is not the same as the ordering of medians, the data have a high level of dispersion and the analysis on group means might be misleading.

I split the distribution of stocks in each visa group into ten quantiles and compare the means within each quantile. Table 4 presents the estimates. In quantiles two through nine the mean in the visa partially exempt category is higher than the mean in the visa required category. Only in quantiles one and ten is this trend reversed. This implies that the ranking of group means might be driven by heavy outliers in quantile ten and by zeros in quantile one. High values in quantile ten artificially increase the mean of the

[7] According to official language data, Russia–Ukraine, Czech Republic–Slovakia, Turkey–Azerbaijan do not share the same official language, but their residents easily understand each other. In many countries of the former Soviet Union, Russian is the language of everyday communication for many people, even though it is an official language only in the Russian Federation.

Table 3: Basic descriptive statistics on the stock of migrants by visa category and year.

	Panel one, 2000			Panel two, 2010			Panel three, pooled		
	vp=0	vp=1	vp=2	vp=0	vp=1	vp=2	vp=0	vp=1	vp=2
				In levels, $stock_{ijt}$					
Mean	11993	8265	58152	17185	9522	36082	14437	8933	43265
St. dev.	140936	49757	274900	184468	63060	189261	162898	57211	220974
Median	64	181	2998	104	160	1329	80	170	1679
Obs.	8217	4156	580	7308	4711	1202	15525	8867	1782
				In logarithms, $\ln(stock_{ijt})$					
Mean	4.78	5.51	7.95	5.11	5.50	7.16	4.94	5.51	7.42
St. dev.	3.09	2.88	2.76	3.18	2.93	2.94	3.14	2.91	2.91
Median	4.55	5.46	8.09	4.91	5.47	7.22	4.71	5.46	7.57
Obs.	7554	3853	557	6909	4322	1177	14463	8175	1734

Table 4: t-test on the differences of means, pooled sample. $\bar{x}$ is the mean of $stock_{ijt}$ and $\bar{y}$ denotes the mean of $\ln(stock_{ijt})$.

	Quantile									
	1	2	3	4	5	6	7	8	9	10
$\bar{x}$(vp=1)-$\bar{x}$(vp=0)	-0.12***	1.59***	8.9***	27.8***	63.8***	130.6***	263.8***	636.4***	808***	-56963.7***
S.E.	(0.02)	(0.07)	(0.19)	(0.44)	(1.01)	(2.24)	(5.82)	(18.42)	(97.36)	(13804)
$\bar{y}$(vp=1)-$\bar{y}$(vp=0)	0.25***	0.67***	0.96***	0.95***	0.81***	0.70***	0.58***	0.50***	0.20***	-0.15***
S.E.	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.05)
$\bar{x}$(vp=2)-$\bar{x}$(vp=0)	5.58***	58.3***	209.5***	497.4***	1100.1***	2440.3***	5067.8***	11196***	29534.3***	238591***
S.E.	(0.50)	(2.29)	(4.77)	(9.16)	(20)	(39.1)	(90.6)	(237.2)	(951.8)	(47068.8)
$\bar{y}$(vp=2)-$\bar{y}$(vp=0)	1.58***	2.72***	3.10***	3.02***	2.87***	2.78***	2.58***	2.37***	1.94***	1.47***
S.E.	(0.09)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.08)

Note: Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

visa required groups and zero values artificially decrease the mean of the visa partially required category. The most prominent examples of country pairs in quantile ten are: migrants in Germany and Israel from the former Soviet Republics, migrants in the United States from Latin America and Chinese migrants in South Korea, Japan and the United States. Typical destination countries with zero values in quantile one are popular tourist destinations, such as: Dominica, Haiti, Ecuador, Maldives, Micronesia and the Cook Islands.

To perform formal testing, I apply Kolmogorov-Smirnov and rank-sum nonparametric tests. At 5% significance level these tests do not reject the hypothesis that the ranking of stocks is $stock_{ij}(vp = 0) < stock_{ij}(vp = 1) < stock_{ij}(vp = 2)$. The logarithmic transformation applied to data reduces dispersion and establishes ranking consistent with the results of the nonparametric tests.

2.6 Econometric model and identification

The baseline specification follows from equations (2.5) and (2.9):

$$y_{ijt} = x'_{ijt}\gamma + d_i + o_j + t_t + \epsilon_{ijt} \quad \forall \;\; i \neq j, \tag{2.11}$$

where $y_{ijt} = \frac{stock_{ijt}}{pop_{it}+pop_{jt}}$, x_{ijt} contains country pair specific covariates (without intercept) and ϵ_{ijt} is a stochastic error which satisfies the Gauss-Markov assumptions. The terms d_i, o_j and t_t capture destination, origin country and time unobserved (latent) heterogeneity. Since these terms are not observed, their effects cannot be estimated. Not accounting for their presence in regression leads to omitted variable bias.

To address this problem one should infer the character of the relationship between d_i, o_j, t_t, and x_{ijt}. The simplest and most restrictive relationship is the mean independence assumption $E[d_i|x_{ijt}] = d_i$, $E[o_j|x_{ijt}] = o_j$ and $E[t_t|x_{ijt}] = t_t$. This assumes that the unobserved heterogeneity is uncorrelated with observed covariates and thus can become part of the error term. Under this assumption regression (2.11) becomes:

$$y_{ijt} = \gamma_0 + x'_{ijt}\gamma + \xi_{ijt}, \tag{2.12}$$

where ξ_{ijt} is a composite error term, $\xi_{ijt} = d_i + o_j + t_t + \epsilon_{ijt}$. Regression (2.12) is a modification of a random effect model and can be consistently estimated by GLS, if its assumptions hold. The Hausman specification test and Wooldridge's test reject the random effect specification in favor of the fixed effect one, implying that the latent heterogeneity is of complex form.

To model the structure of d_i, o_j and t_t I follow three approaches: least squares dummy variable (LSDV), match effects (Mittag, 2012a) and Mundlak's approach (Greene, 2012, Ch. 11; Wooldridge, 2010, Ch. 10). Under the LSDV approach regression (2.11) becomes:

$$y_{ijt} = x'_{ijt}\gamma_0 + D'_i\gamma_1 + O'_j\gamma_2 + T'_t\gamma_3 + \nu_{ijt} \quad \forall \;\; i \neq j, \tag{2.13}$$

where D_i and O_j are column vectors of dummy variables for destination i and origin j. T_t is a column vector of year dummies. In this equation heterogeneity takes the form of group-specific composite intercept $(\gamma_1 + \gamma_2 + \gamma_3)$. Bertoli and Moraga (2013) use origin dummies to control for time invariant characteristics, such as cultural and linguistic proximity.

Under the match effects model individual heterogeneity takes the form of country pair dummy variables. Regression (2.11) can be re-written as:

$$y_{ijt} = x'_{ijt}\gamma_0 + DO'_{ij}\gamma_1 + \nu_{ijt} \quad \forall \;\; i \neq j, \tag{2.14}$$

where DO_{ij} is a column vector of country dyad dummy variables. Since adding $i \times j$ dummies increases the dimensionality of the problem, standard matrix inversion techniques are not practical. Mittag (2012a) develops techniques to address this issue.

Under Mundlak's approach, heterogeneity takes the form of destination, origin, and year group means of all regressors:

$$y_{ijt} = \gamma_0 + x'_{ijt}\gamma_1 + \overline{x}'_{\cdot jt}\gamma_2 + \overline{x}'_{i\cdot t}\gamma_3 + \overline{x}'_{ij\cdot}\gamma_3 + \mu_{ijt} \quad \forall \;\; i \neq j, \tag{2.15}$$

where $\overline{x}'_{\cdot jt} = \frac{1}{N_i}\sum_{i=1}^{N_i} x'_{ijt}$, $\overline{x}'_{i\cdot t} = \frac{1}{N_j}\sum_{j=1}^{N_j} x'_{ijt}$ and $\overline{x}'_{ij\cdot} = \frac{1}{2}\sum_{t=\{2000,\,2010\}} x'_{ijt}$.

The preferred specification is regression (2.15). However, equations (2.13) and (2.14) are also estimated to check for robustness. Since some countries receive migrants from only a few destinations and some countries send migrants to only a few destinations, the intercept is not identified. This leads to the under-identification of the parameters of the visa dummy variables, because they are computed relative to the value of an intercept. This issue persists when the model is checked for robustness using quantile or rolling regressions.

The theoretical model in Section 2.2 predicts one regression in levels (equation 2.5) and the other regression in logarithms (equation 2.9). The PE test (Kmenta, 1990, pp. 521–522) rejects the model in levels in favor of the model in logarithms. Intuitively, the specification in logarithms is preferred because it allows for non-linear relations and smoothes variance thus reducing the amount of outliers and providing a better fit to the data (Wooldridge, 2009, p. 328). For the remainder of the paper I will present estimates of the specification in logarithms only.

2.7 Estimation results

2.7.1 Baseline regression

The OLS estimates of equations (2.13) and (2.15) are reported in Table 5. The exact definitions of variables are given in Table B.2.2. The key dummy variables of interest, vp_1 and vp_2, are positive and significant in all equations. The estimates in specification 1 are biased because they do not account for group specific heterogeneity. In specification 2 this heterogeneity takes the form of destination country, origin country and time dummies.

In specifications 3–5, heterogeneity takes the form of group specific means of variables. Depending on estimated specification, country pairs in groups vp_1 and vp_2 account for around 5–15% more migrants than country pairs with a visa required regime.

Country pairs located further away from each other have fewer migrants. A roughly 1% increase in distance is associated with an 0.2% decrease in the stock of migrants. Country pairs that share a similar language account for about 10% more migration than their counterfactual. Closer cultural and historic links imply slightly more migration. Variables pc_2 and pc_3 are positive and significant in specifications 2–5. This suggests that having been part of the same country, having a common colonizer or sharing a

Table 5: The estimates of equations (2.13) and (2.15).

Variable name	OLS Spec. 1	OLS Spec. 2	OLS Spec. 3	Robust Reg. Spec. 4	LAD Spec. 5
vp_0			base category		
vp_1	0.161***	0.032*	0.139***	0.149***	0.178***
	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)
vp_2	0.642***	0.056*	0.155***	0.139***	0.134***
	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)
ln_dist	-0.374***	-0.217***	-0.190***	-0.132***	-0.112***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
ln_wgap	0.703***	0.509***	-0.058	0.044	-0.030
	(0.06)	(0.18)	(0.07)	(0.06)	(0.05)
ln_stock_prev	0.700***	0.876***	0.908***	0.942***	0.950***
	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)
lang	0.255***	0.164***	0.123***	0.081***	0.063**
	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
pc_1	0.044***	0.017***	0.004	-0.009	-0.003
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
pc_2	0.003	0.055***	0.042***	0.047***	0.057***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
pc_3	-0.182***	0.068***	0.073***	0.069***	0.073***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
cons	-1.959***	-2.727***	-2.627**	0.581	-1.650***
	(0.13)	(0.23)	(1.08)	(0.96)	(0.32)
Adj. R^2	0.733	0.925	0.817	0.860	0.813

Notes: The dependent variable is $\ln(stock_{ijt})$. The number of observations is 18661 in each equation. Spec. 1 does not account for heterogeneity. Spec. 2 includes destination, origin, and year dummies. Spec. 3–5 contain the group means of the variables. In spec. 5 R^2 is computed as the square of the correlation coefficient. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

common border are significant factors in explaining the stocks of migrants. The sign and significance on *ln_stock_prev* variable means strong persistence: destinations with a high diaspora of a certain nationality in the past, will continue to have high stocks of this nationality in the future. These estimates are in line with the findings of other studies mentioned in Table B.2.1.

The sign and significance on the income gap variable, *ln_wgap*, is inconclusive at this stage. Mayda (2010) mentions that the sign might not agree with theoretical predictions because GDP per capita is a measure of average wages and thus ignores variation across skill levels. A worker might get a higher return on skills in a less developed country or move there to start a business. Also, Pedersen et al. (2008) suggest that the effect of income on migration is non-linear. This is addressed in subsection 2.7.2.

Further, I look at whether the policy index can explain variation in the male-female migration gap. The average share of females in the sample is 0.47 with a standard error of 0.001. The hypothesis that the female share in migrant population equals 0.5 is clearly rejected. I regress the share of females on the set of covariates and present the results in Table B.2.6. All four specifications indicate that the share of female migrants in the visa not required group is slightly, but significantly, lower than in the other two groups. The difference is almost 10 more females per 1000 migrants at destination, or approximately 2% (specification 3).

Using the data from Brücker et al. (2013), I compute the share of skilled migrants for each country pair and regress it on the baseline set of covariates. The OLS and alternative estimates of this regression are presented in Table B.2.7. The OLS estimates (specifications 1 and 3) show that the share of skilled migrants does not differ across pairs assigned to different visa categories. However, the estimates from robust regression suggest a negative relationship between these variables. Based on this I conclude that the share of skilled migrants is not larger for country pairs with simplified visa requirements.

2.7.2 Nonlinear effects

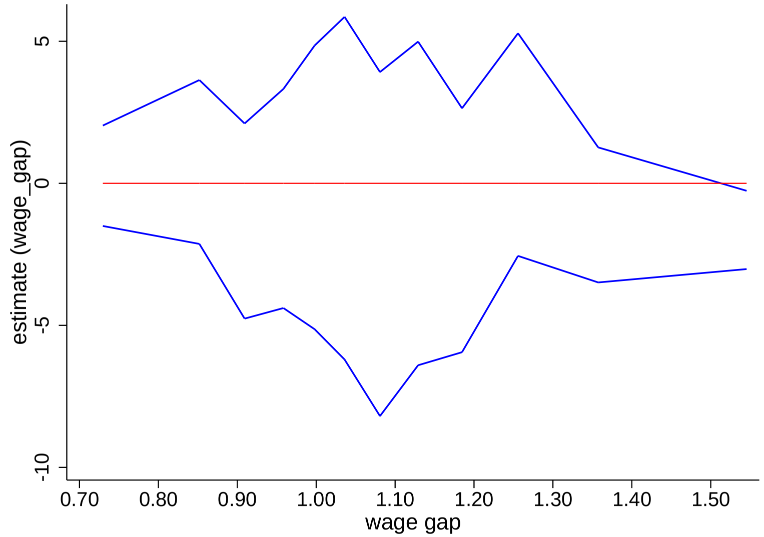
The assignment of country pairs into vp_0, vp_1, and vp_2 groups largely depends on the magnitude of the income gap. Most destinations give visa waivers to origin countries of about the same level of economic development or above. It is difficult to find many instances when an underdeveloped country has a visa-free entry to a developed destination.

The estimate on income gap in the previous section is insignificant in several specifications because, as existing studies suggest, the relationship between income and migration stock is nonlinear. For example, Pedersen et al. (2008) find an inverse U-shaped effect of income at origin on migration. I investigate nonlinearities by income gap quantile and entry visa category. I conjecture that as entry restrictions are lifted, migrants become more responsive to the income gap. In the visa not required category the sign at ln_wgap should not contradict theoretical predictions from Section 2.2 .

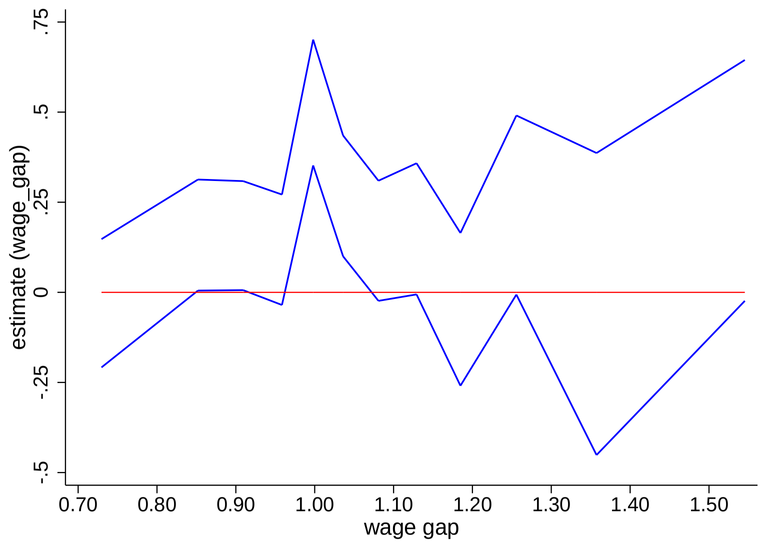
I split the distribution of the income gap into 12 quantiles and estimate equation (2.15) within each quantile and visa category. The confidence intervals of the estimates by quantile are shown in Figure 2 and the estimates by visa group are illustrated in Table B.2.5. Figure 2(a) shows that the wage gap variable does not explain any variation in migrant stock for country pairs in the visa required group. I explain this by the fact that workers do not react to the wage gap due to institutional factors: it is very difficult and costly to obtain an employment visa. When this institutional barrier is partially reduced in Figure 2(b), the income gap affects migration stock positively for country pairs that are not too far off from each other in terms of average income.

Finally, when visa barriers are entirely removed in Figure 2(c), the effect of the wage gap is positive for most of the income gap distribution. For high values of the wage gap the effect reduces to zero for two reasons. First, it is difficult to finance the move when a person comes from a low-income country. The negative effect of poverty on migration is also found in the studies of Mayda (2010) and Pedersen et al. (2008). Second, there is an attrition problem in the data: as the wage gap increases, the visa not required group becomes too small and the effect is not identified.

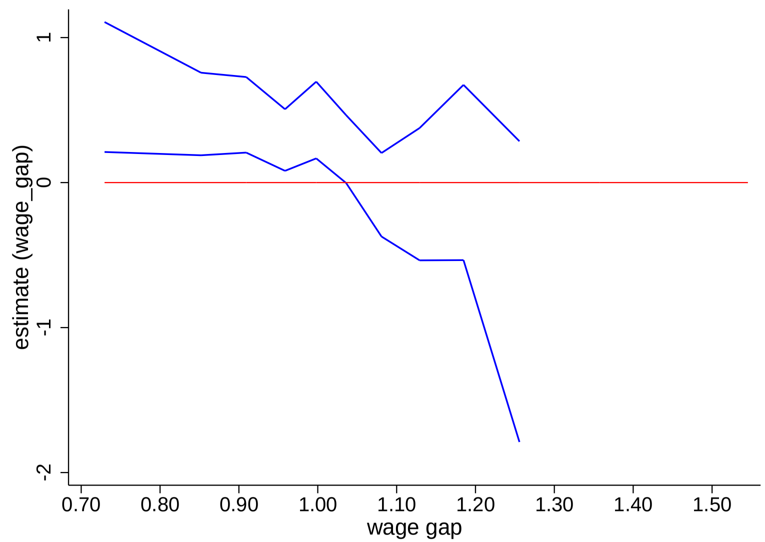
The estimated nonlinear effects are in line with the intuition and findings of other authors. Belot and Ederveen (2012) run a separate regression for country pairs with unrestricted labor mobility. They find that the effects of network, physical distance, cultural and linguistic distance factors weaken or even disappear for this subgroup. The estimation by quantiles is methodologically similar to the study of Beine et al. (2011), who run a rolling regression to estimate the non-linear effects of diaspora size.



(a) *Visa required subsample.*



(b) *Visa partially required subsample.*



(c) *Visa not required subsample.*

Figure 2: 95% confidence intervals of the estimates of wage gap by quantile and visa category.

2.7.3 Difference-in-difference estimates

In this section I discuss the effects of policy changes on the stock of migrants, the share of females and share of skilled migrants. The estimates in levels discussed until now might suffer from one drawback: even though I account for destination country, origin country, and year effects, the visa variables might capture the effects of unobservable characteristics. By construction, the difference-in-difference estimation reduces this kind of bias.

As discussed in Section 2.4, up-shifter country pairs can be of three types: visa required changes to visa partially required (vp_0 to vp_1), visa required changes to visa not required (vp_0 to vp_2) and visa partially required changes to visa not required (vp_1 to vp_2). Symmetrically, down-shifter dyads are: visa partially required changes to visa required (vp_1 to vp_0), visa not required changes to visa required (vp_2 to $vp0$) and visa not required changes to visa partially required (vp_2 to vp_1).

I pool up-shifter pairs in one group and down-shifter pairs in the other group and estimate the effects of policy weakening and tightening. The OLS estimates of the key parameters of interest are illustrated in Table 6. They suggest that before the policy change the up-shifter and down-shifter pairs are not statistically different from non-shifter pairs in terms of migrant stocks and the shares of females and skilled migrants. Controlling for all other covariates, the levels of stocks and the shares of skilled migrants in 2010 are not statistically different from their values in 2000. This, however, is not true for the share of females, which declined in 2010. After the policy change, the up-shifter pairs account for 10% more migrants than their non-shifter counterparts. This is equivalent to 7 more people per 10 mln. of the destination plus origin population. Up-shifter pairs have also smaller shares of females and skilled migrants, both in the magnitude of about 14 migrants per 1000 of migrant stocks at destination. This is equivalent to a 3% decline in the share of females and 3.5% in the share of skilled migrants.

The effect on down-shifters is not symmetric. The estimates of the baseline specification indicate that the introduction of visas is not associated with any significant change in the stocks of migrants, the shares of females and skilled migrants. In Table B.2.8, I provide alternative estimates from robust regression, which suggest an even stronger

Table 6: OLS estimates for policy up-shifters and down-shifters.

	Dep. var. *ln_stock*				Dep. var. *share_female*				Dep. var. *share_skilled*			
	Policy weakening		Policy tightening		Policy weakening		Policy tightening		Policy weakening		Policy tightening	
shift_up	-0.014	-0.009			0.005	0.007**			0.006	0.008		
	(0.03)	(0.04)			(0.00)	(0.00)			(0.01)	(0.01)		
shift_down			0.008	0.036			0.002	0.003			-0.005	-0.004
			(0.04)	(0.07)			(0.00)	(0.00)			(0.01)	(0.01)
y10	0.009	-0.013	0.009	-0.016	-0.008***	-0.008***	-0.009***	-0.008***	-0.001	-0.001	0.006*	0.004
	(0.01)	(0.02)	(0.01)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
shift_up_y10	0.099***	0.102**			-0.013***	-0.014***			-0.014***	-0.014***		
	(0.03)	(0.05)			(0.00)	(0.00)			(0.01)	(0.01)		
shift_down_y10			0.017	0.007			-0.007	-0.007			0.013	0.013
			(0.05)	(0.09)			(0.01)	(0.01)			(0.01)	(0.01)
Obs.	17697	17697	15381	15381	17697	17697	15381	15381	4479	4479	4217	4217
Adj. R^2	0.93	0.82	0.93	0.82	0.68	0.67	0.68	0.67	0.92	0.92	0.86	0.85
Heterogeneity	dummies	means	dummies	means	dummies	means	dummies	means	dummies	means	dummies	means

Notes: All specifications include the baseline covariates. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

result: the tightening of immigration policy is associated with an increase in the share of skilled migrants.

Two major effects generate this asymmetric result: the abolition (introduction) of visas and migrant networks. When immigration policy is weakened, institutional barriers to migration are removed and migrants' costs are reduced. Symmetrically, the tightening of immigration policy imposes additional institutional constraints thus increasing costs. The effect of networks always works in one direction. More migrants of the same nationality or speaking the same language at the destination translates into smaller migration and integration costs, as well as the costs of social exclusion. Numerous studies based on micro and macro data document that migrant communities at the destination support one another in various ways (Munshi, 2003).

When the visa policy weakens, both effects reduce migration costs. However, when visas are imposed, the two effects work in opposite directions. Visas increase costs, but high migrant stocks decrease costs. Shortly after the introduction of visas, the network effect dominates and the stock of migrants does not decrease immediately. With the passing of time, the institutional factor might offset the network effect and the pool of migrants might reduce.

In Figure B.3.1, I plot OLS estimates and associated 95% confidence intervals of the effects of policy changes by the type of change. Figures B.3.1(a) and B.3.1(c) show that the effects of policy weakening on the stocks and the share of skilled migrants are driven by the shifts to the visa not required category (vp_0 to vp_2 and vp_1 to vp_2 in the figures). The effect on the share of females is to a large extent generated by the visa partially required to visa not required shift (vp_1 to vp_2 in the figure).

The effects on down-shifter pairs is not statistically significant for all types of policy tightening changes. In the bottom graph, the effects on v_2 to vp_0 and vp_2 to vp_1 cannot be identified, because the data for these country pairs are not available.

2.7.4 Robustness check

I check the robustness of the results using alternative estimators and estimating the placebo effect. Besides OLS estimates with robust standard errors, in most tables I include alternative results from robust regression or LAD estimates (Verardi and Croux, 2009). If OLS assumptions are not violated, then OLS estimates are preferred. If these assumptions are violated, then alternative estimates should be considered, because they are robust to model misspecifications. To this end, I only present results which are confirmed by OLS and alternative estimates.

I further estimate regression (2.14) with match effects (destination-origin dummies).[8] The advantage of this approach is that it accounts for destination-origin heterogeneity. Since the number of estimated parameters increases, the usual matrix inversion technique is not practical and Mittag (2012a) suggests using a conjugate gradient method. The effects identified on time variant variables suggest a similar picture to the one discussed in sub-section 2.7.1.

In Figure B.3.2, I plot the estimates of the placebo effect obtained from three simulations, 500 iterations each. Within each iteration, I randomly choose a group of country pairs from visa required category and run DiD estimation. Since the visa regime did not change between these country pairs, there should be no statistically significant effect on stocks, their gender and education composition. The DiD estimates in Figure B.3.2 confirm this conjecture. The confidence intervals in Figure B.3.2(a) are relatively wide and contain zero at each iteration. Each of the Figures B.3.2(b) and B.3.2(c) contains about 20 iterations that do not cross the zero line, implying that the placebo effect exists. This, however, is consistent with the definition of a 95% confidence interval: in 5% of cases the estimated interval will not include the value of a true population parameter (type I error). This allows for each simulation to have at most 25 such misclassified iterations.

If the data allowed us to construct the index for a number of years, it would be possible to estimate a dynamic model whereby the stocks of migrants are regressed on different lags of the policy index and other relevant covariates. Using an F-test or an equivalent measure, one would be able to choose an optimal number of lags of the policy index to include in the regression. I expect that the magnitude of the coefficients at the policy variable would decline if the data allowed the inclusion of additional lags.

[8] For the estimation I use $twfe$ STATA module developed by Mittag (2012b).

2.7.5 Discussion of endogeneity

There are three potential sources of endogeneity: omitted variable bias, the inclusion of the lagged dependent variable, and reverse causality. I discuss each source in detail below.

The assignment of country pairs to visa categories is determined by a multiplicity of factors and some of them are not included in the estimation because they are unobserved or unavailable. The created variables vp_0, vp_1, and vp_2 might thus correlate with the error term. To address this problem, each equation is estimated with group specific means of covariates, origin and destination dummies and country pair dummies (match effects). These variables sufficiently account for any unobserved group specific heterogeneity. A similar approach is taken in all studies surveyed in Table B.2.1.

Grogger and Hanson (2011), Belot and Ederveen (2012), Belot and Hatton (2012), Bertoli and Moraga (2013), Mayda (2010), and Pedersen et al. (2008) do not address the issue of the endogeneity of immigration policy *per se*. It is either assumed exogenous or the discussion of endogeneity is not present. Hence, if the endogeneity problem biases my results, these published studies equally share the same problem.[9]

I include a short discussion on four potential instruments for the visa variables: crime rates, visa rejection rates, the shares of refugees, and membership in unions. These instruments are motivated with examples below. However, I am not convinced that any of them satisfies the requirements of a valid instrument as described by Wooldridge (2010).

In 2009, the UK imposed visas on the citizens of South Africa after numerous cases of South African passports being stolen and later misused by other nationalities to get into the UK illegally. Data on crime rates could be a proxy for the frequency of passports stolen. However, data show that per capita crime rates in developing countries are lower that in developed countries, thus invalidating this instrument due to a measurement error.

The visa rejection rate is a proxy for the laxity of visa rules. If the rejection rate is low, then potential migrants do not need enhanced screening and visa rules could be

[9]Mayda (2010) assumes that immigration policy is exogenous (footnote 8 on p. 1253). The remaining studies do sufficiently explain why certain variables proxy for immigration policy. However, they do not address the endogenous nature of the determination of those immigration policy proxy variables.

loosened or entirely abolished. A low rejection rate on B-type visas is one key criterion to joining the US Visa Waiver Program. These data are available only for the US.

In July 2009 Canada temporarily introduced visas for the citizens of the Czech Republic due to a sudden spike in the amount of refugee applications filed on the territory of Canada from Czech visitors. Hence, the number of refugee applications or refugees could be an instrument for changes in immigration policy. However, some destination countries do not separate refugees from economic migrants in the data. The instrument thus does not satisfy the exclusion restriction requirement.

Membership in international organizations or unions can tell something about the credibility of a country in question. After the expansion of the EU in 2004, some of the new member states were added to visa waiver programs in the US, Australia, New Zealand, and South Korea. This information can serve as a regional instrument for the EU origin countries, although it cannot be extended to other continents, where the concept of the EU does not exist.

Mayda (2010), Beine et al. (2011), and Pedersen et al. (2008) address the issue of endogeneity that arises from the inclusion of the lagged dependent variables by using the Arellano and Bond GMM, GEE or IV estimators. My analysis does not suffer from this type of endogeneity for the reason described below. I run the estimation on stocks for 2010 and 2000 and use stocks in 2005 and 1995 to control for network effects. Since my equations are not defined for 2005 and 1995, the Arellano-Bond type correlation does not emerge.

Reverse causality might bias the estimates in a way that push and pull factors affect migration, but migrants also affect wages and the distribution of income. To account for this, I include 5-year averages of the lagged values of GDP per capita and the Gini index. Migrants might also affect the laxity of immigration policy. For example, a destination with more migrants might be more averse to migrants that a destination with fewer migrants. To account for this I use lagged values of immigration policy (March 1998 for 2000 and November 2009 for 2010) and include previous values of the stocks of migrants into the regressions. The approach of including lagged independent variables is quite common in the studies surveyed.

2.7.6 Discussion of results

The abolition of visas reduces institutional barriers to mobility. This pushes more people to migrate in response to cross country income gaps. The increase in migrant stocks from policy affected countries for the period 2000–2010 was around 10%. The growth was mostly in male and unskilled migration. This effect is robust to alternative estimation techniques and specifications.

The selection on gender has to do with the traditional breadwinning role of males in households in many developing economies. Low labor market participation rates for females (World Bank, 2014) combined with low emigration rates (Commander et al., 2013) translate to a gender gap in the migrant stock data. This gap is further widened for country pairs with lax visa rules.

The skill bias is generated by the fact that skilled migrants are less affected by visa restrictions in general. Since many developed destinations have adopted skill-biased immigration policies, it is easier for skilled migrants to obtain a visa. Such migrants are less bound by visa constraints and choose to migrate to destinations that value their skills most.

In contrast, the introduction of visas is not associated with a statistically significant reduction in migrant stocks or their gender or skill composition. Such an asymmetric picture suggests the existence of "immigration policy hysteresis:" it is easy to use immigration policy to increase the stock of migrants, but it is ineffective in reducing migrant stocks in the short run.[10]

Visa restrictions result in being ineffective in the short run for several reasons. First, the introduction of visas affects potential migrants more (migrants in flow data) than current migrants at destination (migrants in stock data). Time is needed for changes in flow data to translate to changes in stocks data. Second, as the estimates in Table 5 illustrate, the magnitude of the network effect exceeds the effect of visas by a factor of six. The immigration policy effect is not strong enough to offset the network effect which is even strengthened after the introduction of visas (see Table B.2.5). It will be interesting to further research if and when the visa effect overtakes the network effect. Unfortunately,

[10] An excellent overview the usage of concept "hysteresis" in Economics is provided by Göcke (2002).

the data at hand do not allow this question to be addressed. Third, firms always demand cheap labor and might lobby for more temporary migrant workers irrespective of visa regime.

2.8 Conclusion

In this paper I achieve two objectives. First, I construct an immigration policy index that varies across sending and receiving countries as well as over time. The index has an intuitive design and clear interpretation in estimation. Second, I use the constructed index to estimate the effects of the introduction and abolition of visas on the stocks of migrants, their gender and education composition.

I find that country pairs with visa partially required and visa not required regimes account for 13% and 15% more migrants respectively than pairs with visa required status. The effect of other determinants also varies by visa category. The effects of wage gap, language and cultural proximity are the strongest in the visa not required category. The effects of distance and diaspora are the strongest in the visa required category.

This result is quite intuitive: if entry visas are required, then migrants move to destinations with large numbers of migrants of their nationality and which are close geographically. If entry visas are absent, then migrants are less restricted in their choice of destination. They usually go to more developed countries and destinations similar in language and culture.

The difference-in-difference estimates show that the introduction of visas is associated with an increase in migrant stocks and a change in gender and skill composition towards more male and less skilled. In contrast, the introduction of visas does not affect the stocks of migrants, their gender or education composition for the period considered. This asymmetric picture hints at the existence of hysteresis effect in how migrant stocks respond to changes in immigration policies.

Bibliography

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., Wacziarg, R., 2003. Fractionalization. Journal of Economic Growth 8 (2), 155–194.

Artuc, E., Docquier, F., Özden, Ç., Parsons, C., 2013. A global assessment of human capital mobility: the role of non-OECD destinations.

Barro, R. J., Lee, J. W., 2013. A new data set of educational attainment in the world, 1950–2010. Journal of Development Economics 104 (0), 184–198.

Beine, M., Docquier, F., Özden, Ç., 2011. Diasporas. Journal of Development Economics 95 (1), 30–41.

Belot, M., Ederveen, S., 2012. Cultural barriers in migration between OECD countries. Journal of Population Economics 25 (3), 1077–1105.

Belot, M. V. K., Hatton, T. J., 2012. Immigrant selection in the oecd. The Scandinavian Journal of Economics 114 (4), 1105–1128.

Bertoli, S., Moraga, J. F.-H., 2013. Multilateral resistance to migration. Journal of Development Economics 102 (0), 79–100.

Brücker, H., Capuano, S., Marfouk, A., 2013. Education, gender and international migration: insights from a panel dataset 1980–2010. mimeo.

Commander, S., Nikolaychuk, O., Vikhrov, D., 2013. Migration from Ukraine: brawn or brain? New survey svidence. IZA Discussion Papers 7348.

Defoort, C., 2008. Long-term trends in international migration: an analysis of the six main receiving countries. Population-E 63 (2), 285–318.

Docquier, F., Rapoport, H., Salomone, S., 2012. Remittances, migrants' education and immigration policy: theory and evidence from bilateral data. Regional Science and Urban Economics 42 (5), 817–828.

Feenstra, R. C., Inklaar, R., M.P., T., 2013. The next generation of the Penn World Table. Available for download at www.ggdc.net/pwt.

Freeman, R. B., Oostendorp, R., 2000. Wages around the world: pay across occupations and countries. Working Paper 8058, National Bureau of Economic Research.

Glaesser, D., Kesler, J., 2013. Tourism visa openness report. Tech. rep., World Tourism Organization, Madrid.

Glick, R., Rose, A. K., 2002. Does a currency union affect trade? The time-series evidence. European Economic Review 46 (7), 1125–1151.

Göcke, M., 2002. Various concepts of hysteresis applied in Economics. Journal of Economic Surveys 16 (2), 167–188.

González, X., Miles-Touya, D., 2014. Admission policies and immigrant skills. Applied Economics Letters 21 (17), 1189–1193.

Greene, W. H., 2012. Econometric analysis, 7th Edition. Prentice hall.

Grogger, J., Hanson, G. H., 2011. Income maximization and the selection and sorting of international migrants. Journal of Development Economics 95 (1), 42–57.

Head, K., Mayer, T., 2015. Chapter 3 – Gravity equations: workhorse, toolkit, and cookbook. In: Elhanan Helpman, K. R., Gopinath, G. (Eds.), Handbook of International Economics. Vol. 4. Elsevier Science, pp. 131–195.

Head, K., Mayer, T., Ries, J., 2010. The erosion of colonial trade linkages after independence. Journal of International Economics 81 (1), 1–14.

Hobolth, M., 2014. Researching mobility barriers: The European visa database. Journal of Ethnic and Migration Studies 40 (3), 424–435.

Hofstede, G., Hofstede, G., Minkov, M., 2010. Cultures and organizations: software of the mind, 3rd Edition. McGraw-Hill Education.

Huddleston, T., Niessen, J., Chaoimh, E. N., White, E., 2011. Migrant integration policy index. Tech. rep., British Council and Migration Policy Group, Brussels.

IATA, 1998. Travel information manual (3).

IATA, 2009. Travel information manual (11).

Inglehart, R., Baker, W. E., 2000. Modernization, cultural change, and the persistence of traditional values. American Sociological Review 65 (1), 19–51.

Kahanec, M., Pytlikova, M., Zimmermann, K. F., 2014. The free movement of workers in an enlarged european union: institutional underpinnings of economic adjustment. IZA Discussion Papers 8456.

Kmenta, J., 1990. Elements of Econometrics, 2nd Edition. Macmillan Publishing.

Lewis, M. P., Simon, G. F., Fennig, C. D., 2013. Ethnologue: languages of the world, 17th Edition. SIL International, Dallas, Texas.

Mayda, A. M., 2010. International migration: a panel data analysis of the determinants of bilateral flows. Journal of Population Economics 23 (4), 1249–1274.

Mayer, T., Zignago, S., 2011. Notes on CEPII's distance measures: the GeoDist database. Working Papers 2011-25, CEPII.

Melitz, J., Toubal, F., 2014. Native language, spoken language, translation and trade. Journal of International Economics 93 (2), 351–363.

Mittag, N., 2012a. New methods to estimate models with large sets of fixed effects with an application to matched employer-employee data from Germany. FDZ-Methodenreport 2012 (2).

Mittag, N., 2012b. TWFE: Stata module to perform regressions with two-way fixed effects or match effects for large datasets. Statistical Software Components, Boston College, Department of Economics.

Munshi, K., 2003. Networks in the modern economy: Mexican migrants in the US labor market. The Quarterly Journal of Economics 118 (2), 549–599.

Neumayer, E., 2010. Visa restrictions and bilateral travel. The Professional Geographer 62 (2), 171–181.

OECD, 2013. International migration outlook 2013. OECD Publishing.

Ortega, F., Peri, G., 2013. The effect of income and immigration policies on international migration. Migration Studies 1 (1), 47–74.

Özden, Ç., Parsons, C., Schiff, M., Walmsley, T., 2011. Where on earth is everybody? World Bank Economic Review 25 (1), 12–56.

Palmer, J., Pytlikova, M., 2013. Labor market laws and intra-European migration: The role of the state in shaping destination choices. CReAM Discussion Paper Series 1311.

Pedersen, P., Pytlikova, M., Smith, N., 2008. Selection and network effects – migration flows into OECD countries 1990–2000. European Economic Review 52 (7), 1160–1186.

Rosenzweig, M., 2010. Global wage inequality and the international flow of migrants. Discussion paper 983, Economic Growth Center, Yale University.

Solt, F., 2014. The standardized world income inequality database, working paper. SWIID Version 5.0, October 2014.

UN, 2011. International migrant stock by origin: the 2010 revision. Population Division, Migration Section.

UN, 2013. Trends in international migrant stock: Migrants by destination and origin.

Verardi, V., Croux, C., 2009. Robust regression in Stata. Stata Journal 9 (3), 439–453.

Vyas, S., Kumaranayake, L., 2006. Constructing socio-economic status indices: how to use principal components analysis. Health Policy and Planning 21 (6), 459–468.

Wooldridge, J. M., 2009. Introductory econometrics: a modern approach, 4th Edition. South-Western Cengage Learning.

Wooldridge, J. M., 2010. Econometrics of cross-section and panel data, 2nd Edition. MIT Press.

World Bank, 2014. World development indicators 2014. Report, World Bank.

Appendix B

B.1 Derivation of Gini Index

The model economy consists of two skill types. Since this is a generalization of the real world with a continuum of skill types, sectoral wages w_L and w_H are not readily available in data. These wages can be derived from an average wage and the Gini index.

An average wage in the model is:

$$W = \alpha \cdot w_H + (1-\alpha) \cdot w_L. \tag{B.16}$$

In Figure B.1.1, I plot the shares of workers against their cumulative wealth.

Unskilled workers contribute $\frac{(1-\alpha)w_L}{W}$ share to total wealth, and skilled workers contribute $\frac{\alpha w_H}{W}$. The slope of the unskilled line OF is $\frac{w_L}{W}$ and the slope of the skilled line is $\frac{w_H}{W}$. The Gini index is the ratio of the area of triangular OCA, S_{OCA}, to the area of triangular OEA, S_{OEA}:

$$G = \frac{S_{OCA}}{S_{OEA}} = 2 \cdot S_{OCA} = \frac{\alpha(1-\alpha)(w_H - w_L)}{(1-\alpha)w_L + \alpha w_H}, \tag{B.17}$$

where:

$$S_{OCA} = \frac{1}{2} - S_{ODC} - S_{DCBE} - S_{CBA}; \quad S_{ODC} = \frac{1}{2}\frac{(1-\alpha)^2 w_L}{(1-\alpha)w_L + \alpha w_H},$$
$$S_{DCBE} = \frac{\alpha(1-\alpha)w_L}{(1-\alpha)w_L + \alpha w_H}; \quad S_{CBA} = \frac{1}{2}\frac{\alpha^2 w_H}{(1-\alpha)w_L + \alpha w_H}.$$

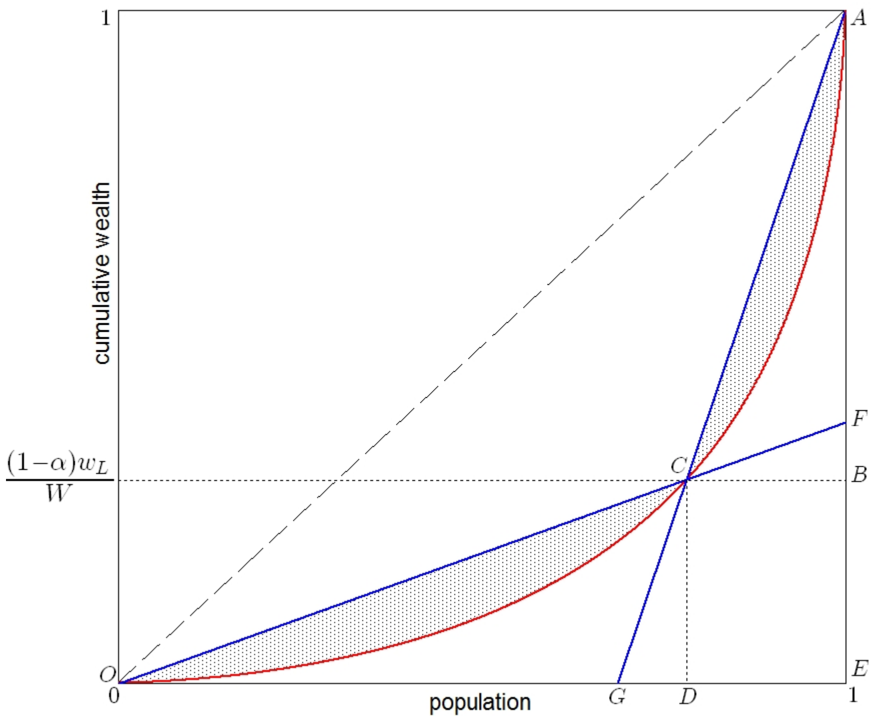


Figure B.1.1: The Gini index for the model.

Using equations (B.16) and (B.17) the sectoral wages are:

$$w_H = W\left(1 + \frac{G}{\alpha}\right), \tag{B.18}$$

$$w_L = W\left(1 - \frac{G}{1-\alpha}\right). \tag{B.19}$$

Equations (B.18) and (B.19) are easy to interpret. If both skill types earned the same wage, it would be equal to the average wage and the Gini index would be zero. Since $w_H > w_L$, the skilled wage is $W\left(\frac{G}{\alpha}\right)$ above the average wage and the unskilled wage is $W\left(\frac{G}{1-\alpha}\right)$ below the average wage.

When compared to the real world with a multiplicity of skill types, the Gini index in equation (B.17) underestimates the degree of inequality. The degree of underestimation equals the shaded area in Figure B.1.1. There are two ways to account for this. One could compute the degree of underestimation (shaded areas) and adjust the Gini index of the model accordingly. Alternatively, one could redefine what a "skilled" worker is, thus adjusting α. I follow the latter approach: an unskilled worker is defined as having zero years of education and a skilled worker has 26 years of education.

B.2 Tables

Table B.2.1: The summary of relevant literature.

	This paper	Belot and Ederveen (2012)	Beine et al. (2011)	Grogger and Hanson (2011)	Bertoli and Moraga (2013)	Mayda (2010)
Data source on stocks/flows	UN (2013), Brücker et al. (2013)	Earlier version of OECD (2013)	Earlier version of Artuc et al. (2013)	Earlier version of Artuc et al. (2013)	Spanish stat. office	Earlier version of OECD (2013)
Time span	2000 and 2010	1990–2003	1990 and 2000	2000	1997–2009, quarterly	1980–1995
Country coverage	All countries	22 OECD countries	30 OECD countries	15 OECD countries	Spain	14 OECD countries
Group heterogeneity	Destination, origin, and year; destination-origin pair	Destination	Destination and origin	Destination-origin pair	Origin-year, quarter	Origin, destination, and year
Immigration policy	Three categories of entry visa restrictions	Dummy for free movement of workers	Share of refugees and dummy for Schengen	Share of refugee visas, dummies for Schengen and visa waiver	Schengen, EU enlargement, Spanish amnesties, visa waiver, various bilateral agreements	Tabulation of changes in legislation at destination
Skill price	Imputed from GDP per capita and the Gini index	GDP per capita	Based on microdata, Rosenzweig (2010)	Based on Freeman and Oostendorp (2000) data and LIS microdata	GDP per capita	GDP per capita
Language proximity	Dummy for overlap between three spoken languages	Own variables based on the proximity of words	Dummy for common language, Head et al. (2010)	Dummy for common language, Head et al. (2010)	Origin fixed effects	Dummy for common lang., Glick and Rose (2002)
Cultural proximity	Principal components based on Head et al. (2010)	Own variables based on Hofstede et al. (2010) and Inglehart and Baker (2000)	Dummy for colonial link after 1945, Head et al. (2010)	Head et al. (2010)	Origin fixed effects	Dummy for colonial relationship, Glick and Rose (2002)
Account for zeros	ln transformation	Zero-inflated negative binomial	Heckman selection	Exclusion of destinations with zeros	Exclusion of destinations with zeros	Tobit model

Table B.2.2: Definitions of covariates and data sources.

Var. name	Definition	Primary source
vp_0, vp_1, vp_2	Dummy variables for visa required, visa partially required and visa not required, respectively.	Own computation using IATA (1998) and IATA (2009) data.
shift_up, shift_down	Dummy variables for country pairs which weakened and tightened their immigration policies, respectively.	Own computation
y10	Dummy variable for 2010.	
shift_up_y10	= shift_up × y10.	Own computation
shift_down_y10	= shift_down × y10.	Own computation
lang	Dummy variable for when two countries share the same or similar language.	Own computation using Lewis et al. (2013) data.
ln_wh_gap, ln_wl_gap	The natural logarithm of imputed wages for skilled and unskilled.	Own computation using Solt (2014), Feenstra et al. (2013) and Barro and Lee (2013) data.
pc_1, pc_2, pc_3	Principal components that describe cultural and historic proximity between a destination and origin country.	Own computation using Head et al. (2010) data.
ln_dist	The natural logarithms of physical distance (in km) between a destination and origin country.	Mayer and Zignago (2011)
ln_wgap	the ratio of the logarithms of GDP per capita (destination) to the GDP per capita (origin).	Feenstra et al. (2013)
ln_stock_prev, ln_female_prev, ln_high_prev	The stocks of total, female, and skilled migrants, respectively, in 1995 and 2005 divided by destination country population.	UN (2013), Brücker et al. (2013)
ln_stock	The stock of migrants divided by destination plus origin county population.	UN (2013)
female_ share	The share of females in total stock.	UN (2013)
skilled_share	The share of skilled in total stock.	Brücker et al. (2013)

Table B.2.3: Examples of county pairs in each visa category in 2010.

Visa required:
USA←(MEX, CHN, IND, KOR, VNM, JAM, DOM),
EEA←(TUR, RUS, KAZ, MKD, MAR, DZA, TUN),
BRA←(JPN, USA, CHN, LBN, EGY, MEX, CUB),
ARG←(CHN, UKR, CUB, SYR, LBN, ARM, LAO, MAR, IRN),
ARG←(DZA, IND, IDN, SAU, PHL, NGA, COG, ETH, TZA),
ZAF←(SLB, TCD, GMB, TGO, BTN, SLV, KAZ, QAT, DJI, SUR),
ISR←(MAR, UKR, ETH, IRQ, VEN, DZA, YEM, TUR, ZMB, NRU),
KOR←(CHN, VNM, PHL, IDN, MNG, UZB, LKA, BGD, NPL),
AUS←(CHN, IND, VNM, PHL, ZAF, LBN, IDN, HRV, THA),
ARE←(IND, BGD, PAK, PHL), IND-(BGD, PAK),
JPN←THA, CAN←(ZAF, MAR, EGY),
GAB←(BEN, CMR).

Visa partially required:
EEA←(USA, CAN, BRA, KOR, ISR),
EEA←(MEX, AUS, NZL, CHL, MYS, VEN),
USA←(most EEA); AUS←(most EEA),
MYS←(IDN, BGD, NZL, IRQ, GTM, LBN),
ZAF←(MOZ, ZWE, LSO, GBR, NAM, SWZ, MLW, ZWE),
ARG←(BRA, most EEA, RUS, COL, MEX, TUR, ISR, PAN, ZAF),
CHL←(PER, ARG, BOL, ECU, COL, USA, most EEA, TUR).

Visa not required:
USA↔CAN, AUS↔NZL,
migration within the EEA,
migration within most of the former Soviet Union,
migration between (Bahrain, Kuwait, Oman,
Qatar, Saudi Arabia, and United Arab Emirates),
IND←(NPL, BTN); between (DZA-MAR-TUN),
LBR←(NER, BFA, CIV, GNB, TGO, SEN),
between (UGA, ERI, KEN).

Note: CZE←UKR means migration to the Czech Republic from Ukraine.

Table B.2.4: Examples of county pairs by allowed durations of visa free stay as of 2010.

Duration of visa free stay	Example of country pairs
$[3, 30)$ days	MYS←CHN, VNM←BRN, DMA←CUB, LSO←ZWE, BTN←IND.
30 and 31 days	EGY←(ARG, AUS), ARE←AUS, IDN←CHN.
$[45, 90]$ days	EU←(AND, ARG, AUS, BRA, BRB, USA, CAN), (ISR, TUR)←EU, URY←CHL, ZAF←ZWE.
$[120, 365]$ days	CAN←(AUS, BEL, BRB), MEX←(most EU, CHL), GBR←(ISR, MUS), GEO←(most countries).

Table B.2.5: Estimates by visa category.

Var. name	$vp_0 = 1$ Spec. 1	$vp_2 = 1$ Spec. 2	$vp_2 = 1$ Spec. 3
ln_dist	-0.256***	-0.062***	-0.052
	(0.02)	(0.02)	(0.04)
ln_wgap	0.105	-0.083	1.766***
	(0.07)	(0.13)	(0.36)
ln_stock_prev	0.935***	0.952***	0.836***
	(0.01)	(0.01)	(0.02)
lang	0.092**	0.003	0.197**
	(0.04)	(0.05)	(0.08)
pc_1	-0.011	0.001	-0.021
	(0.01)	(0.01)	(0.02)
pc_2	0.022	0.064***	0.082***
	(0.02)	(0.01)	(0.02)
pc_3	0.033*	0.074***	0.054**
	(0.02)	(0.02)	(0.03)
cons	1.319	-1.419	-1.640
	(1.27)	(1.62)	(3.26)
Obs.	10672	6555	1434
Adj. R^2	0.859	0.836	0.838

Notes: The dependent variable is $\ln(stock_{ijt})$. Specifications 1, 2, and 3 are estimated on the subsamples of country pairs with visa required, visa partially required, and visa not required regimes, respectively. All specifications include group specific means of variables. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table B.2.6: The estimates of the female share on the set of covariates.

Variable name	OLS Spec. 1	Robust Reg. Spec. 2	OLS Spec. 3	Robust Reg. Spec. 4
vp_0		base category		
vp_1	-0.001	-0.001**	0.001	-0.002***
	(0.00)	(0.00)	(0.00)	(0.00)
vp_2	-0.012***	-0.006***	-0.009***	-0.007***
	(0.00)	(0.00)	(0.00)	(0.00)
ln_dist	0.003**	0.002***	0.003***	0.003***
	(0.00)	(0.00)	(0.00)	(0.00)
ln_wgap	-0.032	-0.006	0.015**	0.007***
	(0.02)	(0.01)	(0.01)	(0.00)
fratio_prev	0.704***	0.983***	0.710***	0.983***
	(0.01)	(0.00)	(0.01)	(0.00)
ln_female_prev	0.000	0.001***	0.001*	0.001***
	(0.00)	(0.00)	(0.00)	(0.00)
lang	0.006**	0.000	0.002	-0.001
	(0.00)	(0.00)	(0.00)	(0.00)
pc_1	0.001*	0.000	0.000	0.000
	(0.00)	(0.00)	(0.00)	(0.00)
pc_2	0.002***	0.000	0.002***	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)
pc_3	-0.001	-0.000	-0.001	-0.001*
	(0.00)	(0.00)	(0.00)	(0.00)
cons	0.16***	-0.016	-0.62***	-0.363***
	(0.04)	(0.02)	(0.1)	(0.03)
Adj. R^2	0.681	0.979	0.663	0.972

Notes: The dependent variable is the share of females in stocks. Spec. 1 and 2 include destination, origin, and year dummies. Spec. 3 and 4 contain the group means of the variables. The number of observations is 18661 in each equation. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table B.2.7: The estimates of the share of skilled migrants on the set of covariates.

	OLS Spec. 1	Rob. Reg Spec. 2	OLS Spec. 3	Rob Reg. Spec. 4	LAD Spec. 5
vp_0			base category		
vp_1	0.003	-0.003	0.002	-0.003	-0.005
	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
vp_2	0.010	-0.015***	0.010	-0.013***	-0.018***
	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)
ln_dist	-0.004*	-0.006***	-0.003	-0.005***	-0.004**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ln_high_prev	-0.013***	-0.004***	-0.010***	-0.004***	-0.005***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
share_high_prev	0.757***	0.954***	0.755***	0.932***	0.918***
	(0.02)	(0.00)	(0.02)	(0.01)	(0.01)
ln_wh_gap	-0.003	0.006	0.000	0.002*	0.002
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
ln_wl_gap	-0.002	-0.002	-0.003*	-0.003**	-0.002**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
lang	0.016***	0.005**	0.012***	0.002	0.002
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
pc_1	0.000	0.000	-0.001	0.000	0.001*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
pc_2	-0.001*	-0.000	-0.001*	-0.001	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
pc_3	0.002*	-0.000	0.002**	-0.000	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
cons	-0.058	-0.030	0.356***	-0.067	-0.248***
	(0.04)	(0.02)	(0.13)	(0.07)	(0.08)
Adj. R^2	0.854	0.962	0.842	0.944	0.83

Notes: The dependent variable is the share of skilled migrants in stocks. Spec. 1 and 2 include destination, origin, and year dummies. Spec. 3–5 have the group means of the variables. In spec. 5 R^2 is computed as the square of the correlation coefficient. The number of observations is 4747 in each equation. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table B.2.8: Robust regression estimates of the model in Table 6.

	Dep. var. *ln_stock*				Dep. var. *share_female*				Dep. var. *share_skilled*			
	Policy weakening		Policy tightening		Policy weakening		Policy tightening		Policy weakening		Policy tightening	
shift_up	0.064***	-0.006			0.001	0.002*			0.009**	0.012***		
	(0.02)	(0.03)			(0.00)	(0.00)			(0.00)	(0.00)		
shift_down			0.048	0.064			0.003**	0.003			-0.018***	-0.011*
			(0.03)	(0.06)			(0.00)	(0.00)			(0.01)	(0.01)
y10	0.041***	0.038**	0.042***	0.035*	-0.005***	-0.005***	-0.005***	-0.005***	-0.009***	-0.007***	-0.009***	-0.004**
	(0.01)	(0.02)	(0.01)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
shift_up_y10	-0.012	0.091**			-0.005***	-0.006***			-0.020***	-0.020***		
	(0.02)	(0.04)			(0.00)	(0.00)			(0.00)	(0.00)		
shift_down_y10			-0.026	-0.004			-0.005**	-0.004			0.028***	0.021**
			(0.04)	(0.08)			(0.00)	(0.00)			(0.01)	(0.01)
Obs.	17697	17697	15381	15381	17697	17697	15381	15381	4479	4479	4217	4217
Adj. R^2	0.93	0.82	0.93	0.82	0.68	0.67	0.68	0.67	0.92	0.92	0.86	0.85
Heterogeneity	dummies	means	dummies	means	dummies	means	dummies	means	dummies	means	dummies	means

Notes: All specifications include the baseline covariates. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table B.2.9: *Policy changes 2000–2010. A dot means no change.*

Country	ISO code	d(i,vp=0)	d(i,vp=1)	d(i,vp=2)
Jamaica	JAM	-0.01	-0.21	0.22
Andorra	AND	-0.33	0.17	0.15
Holy See	VAT	-0.05	-0.10	0.15
San Marino	SMR	-0.05	-0.10	0.15
Latvia	LVA	-0.08	-0.06	0.15
Lithuania	LTU	-0.15	.	0.15
Cyprus	CYP	0.07	-0.22	0.15
Belgium	BEL	-0.05	-0.09	0.15
Estonia	EST	-0.13	-0.02	0.15
Romania	ROU	-0.21	0.06	0.15
Poland	POL	-0.11	-0.04	0.15
Italy	ITA	-0.05	-0.09	0.15
Hungary	HUN	-0.03	-0.11	0.15
Malta	MLT	0.28	-0.43	0.15
Slovenia	SVN	-0.08	-0.07	0.15
Iceland	ISL	-0.01	-0.14	0.15
Switzerland	CHE	0.05	-0.20	0.14
Czech Republic	CZE	-0.11	-0.03	0.14
Liechtenstein	LIE	0.05	-0.20	0.14
Slovakia	SVK	-0.11	-0.03	0.14
Guinea	GIN	0.00	-0.10	0.10
Trinidad and Tobago	TTO	-0.18	0.08	0.10
United Kingdom	GBR	-0.01	-0.06	0.07
Monaco	MCO	-0.10	0.03	0.07
Greece	GRC	-0.08	0.01	0.07
Germany	DEU	-0.03	-0.03	0.07
Sweden	SWE	0.05	-0.12	0.07
Ireland	IRL	-0.08	0.01	0.07
Denmark	DNK	-0.04	-0.03	0.07
Austria	AUT	-0.05	-0.01	0.07
Luxembourg	LUX	-0.05	-0.02	0.07
Norway	NOR	0.06	-0.13	0.07
Spain	ESP	-0.05	-0.02	0.07
Finland	FIN	0.02	-0.09	0.07
France	FRA	-0.10	0.03	0.07
Portugal	PRT	-0.10	0.03	0.07
Netherlands	NLD	-0.05	-0.01	0.06
Tunisia	TUN	0.02	-0.04	0.02
Northern Mariana Islands	MNP	0.83	-0.85	0.02
Kyrgyzstan	KGZ	-0.15	0.13	0.01
Liberia	LBR	-0.01	.	0.01
Eritrea	ERI	-0.01	.	0.01
Marshall Islands	MHL	-0.20	0.19	0.01
Angola	AGO	-0.01	.	0.01
Micronesia (Federated States of)	FSM	.	-0.01	0.01
India	IND	.	-0.01	0.01
Ethiopia	ETH	-0.18	0.17	0.00
Burundi	BDI	-0.01	0.00	0.00
American Samoa	ASM	0.84	-0.85	0.00
Jordan	JOR	-0.03	0.03	.
Chile	CHL	-0.06	0.06	.
The former Yugoslav Republic of Macedonia	MKD	-0.05	0.05	.
Barbados	BRB	0.01	-0.01	.
Grenada	GRD	-0.02	0.02	.
Bahrain	BHR	-0.07	0.07	.
Croatia	HRV	-0.11	0.11	.
Saudi Arabia	SAU	.	.	.
Zambia	ZMB	0.03	-0.03	.
Albania	ALB	-0.10	0.10	.
Zimbabwe	ZWE	0.25	-0.25	.
Bermuda	BMU	-0.73	0.73	.
Lebanon	LBN	-0.37	0.37	.
Malawi	MWI	-0.04	0.04	.
Bosnia and Herzegovina	BIH	-0.12	0.12	.
Namibia	NAM	-0.07	0.07	.
Cuba	CUB	-0.01	0.01	.
Honduras	HND	-0.22	0.22	.

continued on next page

Table B.2.9 – continued from previous page

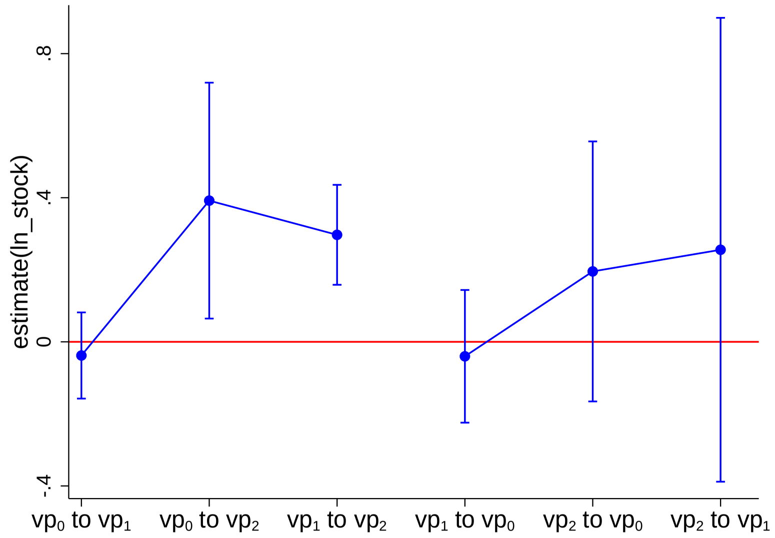
Country name	ISO code	d(i,vp=0)	d(i,vp=1)	d(i,vp=2)
Somalia	SOM	0.00	0.00	.
Paraguay	PRY	-0.13	0.13	.
Aruba	ABW	-0.24	0.24	.
Iraq	IRQ	-0.15	0.15	.
Bolivia (Plurinational State of)	BOL	-0.14	0.14	.
United States Virgin Islands	VIR	0.00	0.00	.
Burkina Faso	BFA	0.19	-0.19	.
Mali	MLI	-0.90	0.90	.
Tuvalu	TUV	-0.57	0.57	.
Uganda	UGA	-1.00	1.00	.
Togo	TGO	.	.	.
Côte d'Ivoire	CIV	0.01	-0.01	.
Seychelles	SYC	0.00	0.00	.
New Zealand	NZL	-0.10	0.10	.
Indonesia	IDN	-0.14	0.14	.
Equatorial Guinea	GNQ	.	.	.
Colombia	COL	-0.28	0.28	.
Mauritius	MUS	-0.18	0.18	.
Bangladesh	BGD	-0.73	0.73	.
Republic of Korea	KOR	-0.28	0.28	.
Lao People's Democratic Republic	LAO	-0.05	0.05	.
Iran (Islamic Republic of)	IRN	0.00	0.00	.
Guatemala	GTM	-0.25	0.25	.
Mauritania	MRT	0.01	-0.01	.
Madagascar	MDG	.	.	.
Kuwait	KWT	-0.19	0.19	.
China, Hong Kong Special Administrative Region	HKG	0.01	-0.01	.
United Arab Emirates	ARE	-0.16	0.16	.
Gabon	GAB	0.03	-0.03	.
Peru	PER	-0.08	0.08	.
Chad	TCD	.	.	.
Dominica	DMA	-0.87	0.87	.
French Guiana	GUF	-0.09	0.09	.
Kiribati	KIR	-0.11	0.11	.
Oman	OMN	-0.33	0.33	.
Serbia	SRB	-0.17	0.17	.
Thailand	THA	-0.05	0.05	.
Malaysia	MYS	0.01	-0.01	.
French Polynesia	PYF	-0.09	0.09	.
Djibouti	DJI	-0.90	0.90	.
Cape Verde	CPV	-0.01	0.01	.
Argentina	ARG	-0.12	0.12	.
Senegal	SEN	.	.	.
Nicaragua	NIC	-0.89	0.89	.
Réunion	REU	-0.08	0.08	.
Fiji	FJI	-0.10	0.10	.
Saint Vincent and the Grenadines	VCT	0.04	-0.04	.
Cambodia	KHM	.	.	.
Samoa	WSM	.	.	.
Nigeria	NGA	0.00	0.00	.
Niue	NIU	0.85	-0.85	.
Australia	AUS	-0.10	0.10	.
Philippines	PHL	0.03	-0.03	.
Palau	PLW	.	.	.
El Salvador	SLV	-0.27	0.27	.
Lesotho	LSO	0.04	-0.04	.
Mozambique	MOZ	-1.00	1.00	.
Montenegro	MNE	-0.21	0.21	.
Bahamas	BHS	-0.09	0.09	.
Japan	JPN	-0.05	0.05	.
Kenya	KEN	-0.68	0.68	.
Papua New Guinea	PNG	0.41	-0.41	.
United Republic of Tanzania	TZA	-0.47	0.47	.
Mexico	MEX	-0.11	0.11	.
Maldives	MDV	.	.	.
Benin	BEN	0.00	0.00	.
Mayotte	MYT	-0.09	0.09	.
Bhutan	BTN	0.00	0.00	.
Turkey	TUR	-0.06	0.06	.

continued on next page

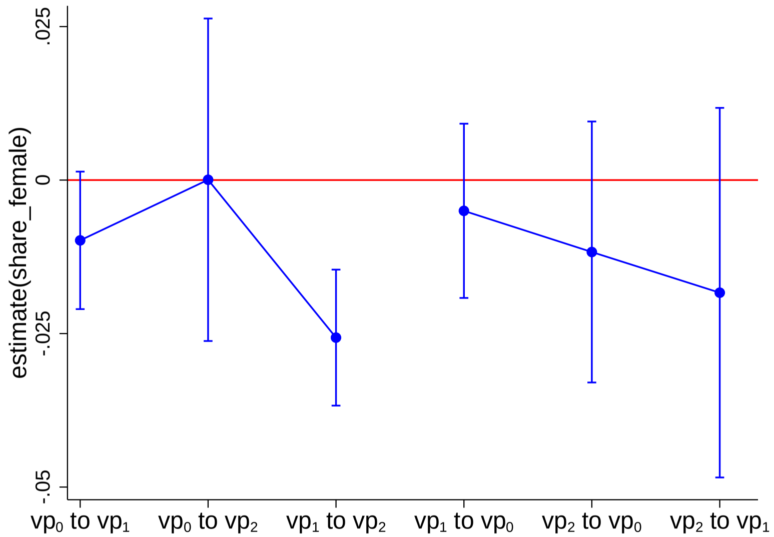
Table B.2.9 – continued from previous page

Country name	ISO code	d(i,vp=0)	d(i,vp=1)	d(i,vp=2)
Guadeloupe	GLP	-0.09	0.09	.
Sao Tome and Principe	STP	1.00	-1.00	.
Qatar	QAT	-0.16	0.16	.
Dominican Republic	DOM	-0.18	0.18	.
Tonga	TON	0.63	-0.63	.
Ecuador	ECU	-0.07	0.07	.
Gambia	GMB	0.00	0.00	.
Comoros	COM	.	.	.
Panama	PAN	-0.05	0.05	.
Turks and Caicos Islands	TCA	-0.06	0.06	.
Morocco	MAR	-0.08	0.08	.
Botswana	BWA	-0.17	0.17	.
Nauru	NRU	-0.19	0.19	.
Saint Kitts and Nevis	KNA	0.33	-0.33	.
Vietnam	VNM	-0.07	0.07	.
Brazil	BRA	-0.11	0.11	.
Brunei Darussalam	BRN	-0.14	0.14	.
Guyana	GUY	-0.02	0.02	.
Sri Lanka	LKA	-0.04	0.04	.
Algeria	DZA	0.03	-0.03	.
Guam	GUM	-0.03	0.03	.
Israel	ISR	-0.09	0.09	.
Myanmar	MMR	.	.	.
United States of America	USA	0.00	0.00	.
Nepal	NPL	0.05	-0.05	.
Haiti	HTI	-0.87	0.87	.
China	CHN	-0.02	0.02	.
Ghana	GHA	0.01	-0.01	.
Vanuatu	VUT	-0.06	0.06	.
Rwanda	RWA	-0.04	0.04	.
New Caledonia	NCL	-0.06	0.06	.
Puerto Rico	PRI	0.00	0.00	.
Antigua and Barbuda	ATG	0.02	-0.02	.
Sierra Leone	SLE	0.00	0.00	.
Egypt	EGY	-0.05	0.05	.
Uruguay	URY	-0.09	0.09	.
Costa Rica	CRI	-0.08	0.08	.
Cook Islands	COK	.	.	.
Canada	CAN	0.02	-0.02	.
Bulgaria	BGR	-0.08	0.08	.
Cameroon	CMR	.	.	.
Pakistan	PAK	.	.	.
Saint Lucia	LCA	-0.13	0.13	0.00
Solomon Islands	SLB	0.29	-0.28	0.00
Belize	BLZ	-0.08	0.09	0.00
Venezuela (Bolivarian Republic of)	VEN	-0.13	0.13	0.00
Syrian Arab Republic	SYR	.	0.00	0.00
Ukraine	UKR	-0.13	0.14	-0.01
Sudan	SDN	-0.01	0.02	-0.01
Russian Federation	RUS	0.02	-0.01	-0.01
Azerbaijan	AZE	0.00	0.00	-0.01
Uzbekistan	UZB	0.01	0.00	-0.01
Mongolia	MNG	-0.01	0.03	-0.02
Singapore	SGP	0.02	0.00	-0.02
Gibraltar	GIB	0.03	.	-0.03
Swaziland	SWZ	-0.10	0.14	-0.04
Kazakhstan	KAZ	0.03	0.01	-0.04
Belarus	BLR	0.02	0.02	-0.05
Central African Republic	CAF	0.06	.	-0.06
Turkmenistan	TKM	0.07	.	-0.07
Libya	LBY	0.06	0.01	-0.07
Yemen	YEM	-0.13	0.22	-0.09
Suriname	SUR	-0.01	0.12	-0.11
South Africa	ZAF	-0.03	0.15	-0.12
China, Macao Special Administrative Region	MAC	-0.16	0.33	-0.17
Montserrat	MSR	-0.14	0.45	-0.31
Cayman Islands	CYM	0.00	0.43	-0.44

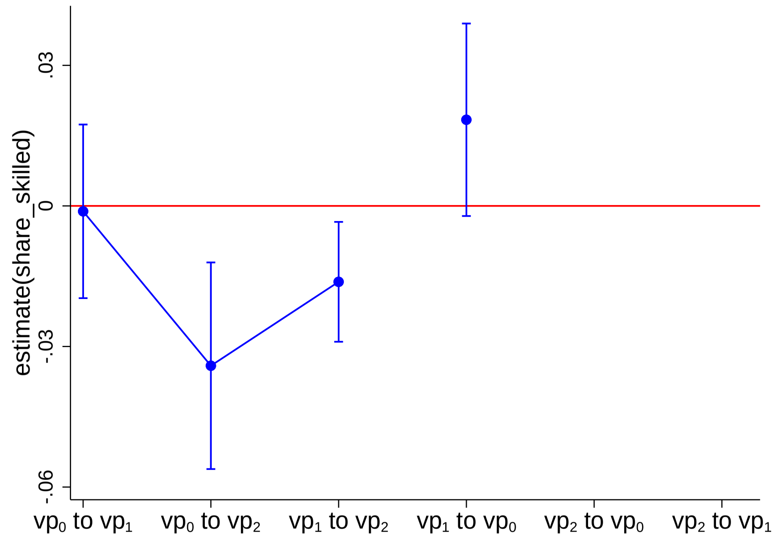
B.3 Figures



(a) *Stocks of migrants equation.*

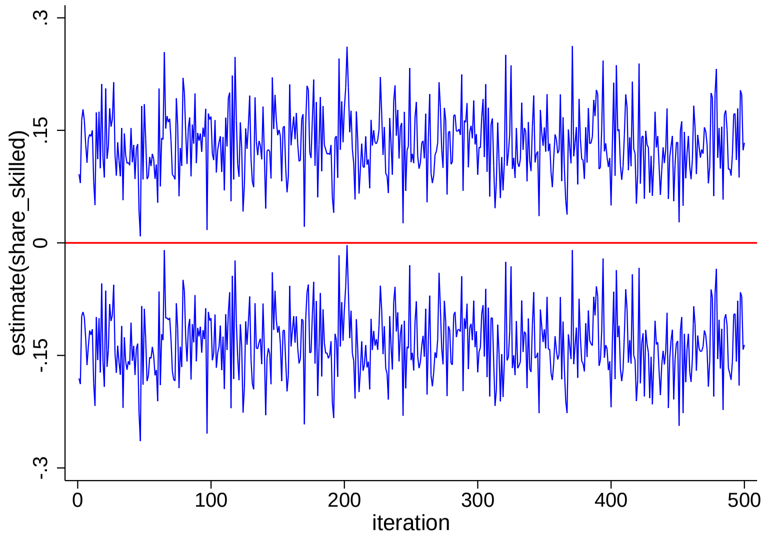


(b) *Share of female migrants equation.*

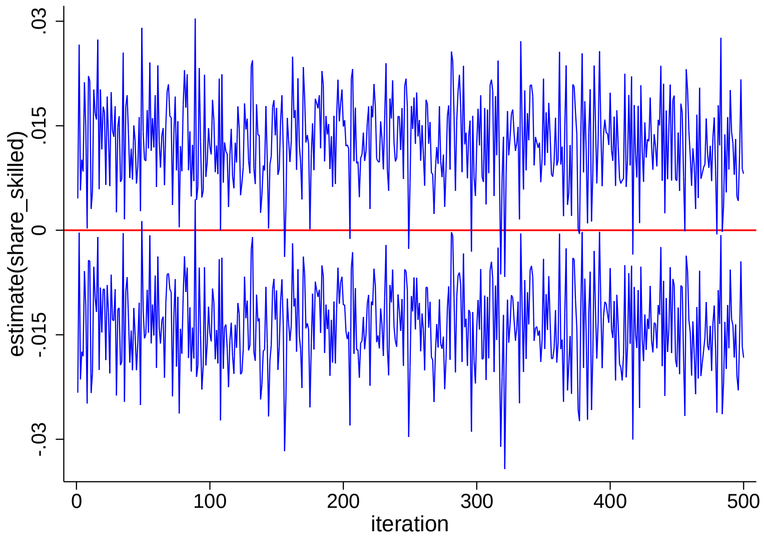


(c) *Share of skilled migrants equation.*

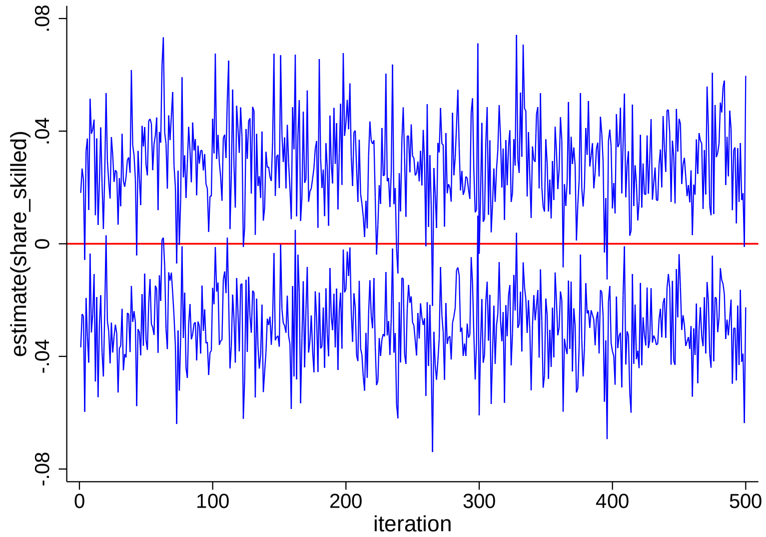
Figure B.3.1: DiD estimates by the type of policy change.



(a) *Stocks of migrants regression.*



(b) *Share of female migrants regression.*



(c) *Share of skilled migrants regression.*

Figure B.3.2: The estimates of the placebo effect.

Chapter 3

Migration from Ukraine: Brawn or Brains? New Survey Evidence

Abstract

We study selection and labor market outcomes among Ukrainian migrants using unique survey data collected in Ukraine in August–October 2011. We find that migrants are positively selected on age, education and pre-migration income. However, this is not associated, as might be expected, with their labor market outcomes. Notably, around half of the migrants are employed in occupations for which they are overeducated and 20% in occupations less skill demanding than their pre-migration match. This phenomenon can be explained by the absence of the conventional link between education and skills in Ukraine and poor cross-border transferability of human capital obtained in Ukraine. We combine the decision to emigrate and downshift into the unified framework of bivariate probit, which we augment to account for unobserved individual heterogeneity in labor market achievements. Further, we estimate three types of network effects on the family level and analyse gender differences.

Keywords: emigration; selection; overeducation; occupation downshift; individual heterogeneity; bivariate probit; survey data

JEL classification: F22; J24; O15; R23

This paper is a joint work with Simon Commander and Olexandr Nikolaychuk. The paper is Revise and Resubmit to the Journal of Comparative Economics at the time of the submission of the thesis. The earlier version was published as IZA discussion paper No. 7348. We are thankful to the participants of seminars/workshops at EBRD (2012), Czech Economic Society (2012), Parthenope University in Naples (2014) and IOS in Tutzing (2014) for useful discussions. The research was generously supported by the European Bank for Reconstruction and Development.

3.1 Introduction

In recent decades, Ukraine has seen a significant increase in cross-border migration, as well as a diversification in the direction of that migration away from other former Soviet states. It is therefore interesting to know the characteristics of those migrants because the welfare change from emigration is to a large extent determined by those characteristics.

In this paper we research the patterns of selection of Ukrainian migrants, i.e. Ukraine nationals living outside Ukraine, and investigate their labor market outcomes. For this purpose we conduct an ad hoc survey, which is different, and in many respects richer, than existing studies in that it asks an exhaustive set of individual and household level questions and has a relatively large sample size.[1] The survey is designed to test for self-selection into emigration, return migration, and prospective emigration based on key observable characteristics.

Since we track migrants' employment details before and after emigration, we are able to identify the process of occupation downshift, whereby migrants take jobs with characteristics that do not match well their capabilities. To the best of our knowledge this feature has rarely been identified in empirical studies of emigration from Ukraine. The paper also provides the first detailed look at the properties of migration from this large emerging market in the aftermath of the financial and economic crises.

The collected data show that the emigration rate is around 10% and the migrants are positively selected in terms of age, education, and pre-migration income. The average wage gap across comparable occupations in Ukraine and outside Ukraine is more than of the order of two. This plays a significant pull factor in the individual decision to emigrate. There is a large variation across destinations. However, Russia is the largest recipient of migrants from most regions of Ukraine.

Migrants' occupation choices abroad often do not correspond with their observable education levels. Under the assumption that skills and education are tightly correlated, this suggests that migrants tend to downshift when finding work abroad. Nearly half of

[1]We surveyed in the source country thus avoiding the issues of unknown distribution of migrant population in destination countries. This also enables us to capture individuals who will have various statuses in destination countries and are thus difficult to find, for example irregular migrants. By design, the data collected can be compared to only three existing studies: Mahmoud and Trebesch (2010); Libanova (2009) and ETF (2008). The face-to-face interviews were carried out by a professional marketing company called "UMP."

the migrants in our survey data are employed in occupations for which they are clearly overqualified. This is, of course, a feature found in some other studies of migrants' occupational choices (Mattoo et al., 2008 and Kostenko et al., 2012, to mention a few studies). Occupation downshift, also referred to in the literature as "brain waste", "overeducation", "overqualification" or "underemployment", persists across all education categories irrespective of the destination country. In this paper, we look at the determinants of emigration and occupation downshift and investigate them in a unified framework.

The downshift phenomenon can be explained by a combination of factors: job search costs, language skills, institutional barriers (visa, employment permits, qualification exams), and the degree of the transferability of human capital obtained in Ukraine. Part of this may also be motivated by large income gaps between better paid and more skilled occupations in Ukraine and low paid occupations abroad. The gap is in favor of the latter. This implies that there might still be incentives for an individual with relatively low migration costs to downshift or choose an occupation that is seemingly a bad match for their educational background. We also find that many Ukrainians downshift already in Ukraine, which points to the role of individual unobservables and a wedge between skills required by the market and skills acquired at universities.

Our paper is one of the few studies which systematically analyzes emigration and labor market outcomes from a large source country such as Ukraine. This study bears particular relevance in the light of the ongoing transition in Ukraine.

3.2 Literature review

This study is related to the body of research on self-selection which investigates how the characteristics of migrants differ from non-migrants and the destination country population (Borjas, 1987; Chiquiar and Hanson, 2005; Moraga, 2011, among others). These characteristics define how migrants affect the receiving and sending economies, thus linking our study also to the literature on the labor market outcomes of migrants and the brain drain / brain gain literature (see, for example, Gibson and McKenzie, 2012).

Borjas (1987) first applied the self-selection framework to study the quality of migrants in the US. He defined three types of selection: positive selection (migrants are above average in income distribution, and thus unobservable skills or abilities, in both sending and receiving countries), selection of refugees (migrants are above average only in the receiving country) and negative selection (migrants are below average in both countries). Using data from the 1970 and 1980 population censuses, he found evidence for positive selection and an increase in earnings over time for migrants from Western Europe and negative selection and a decrease in earnings over time for those from less developed countries.

Further empirical evidence on selection has been mixed. In the context of Mexico-US migration Chiquiar and Hanson (2005) and McKenzie and Rapoport (2010) find that the probability of emigration increases for those in the middle and high sections of the education distribution (positive selection). Using longitudinal data, Moraga (2011) finds negative selection of migrants and shows that the distribution of the would-be-migrants' earnings dropped in the pre-migration quarter. Elsewhere, Rooth and Saarela (2007) find that Finnish migrants in Sweden during 1989–1990 had on average 1 year less of schooling than non-migrants. This is consistent with the fact that for a decade prior to the period considered Sweden had a smaller return to observable skills than Finland.

In part due to data availability, the vast majority of studies research self-selection from the perspective of receiving countries. This approach has its own problems. These include the lack of representativeness of a survey sample in terms of certain migration categories (irregular migrants, for example). Hanson (2006) estimates that the undercount of illegal migrants in the stock data in the US Census is around 10–25%. Moraga (2011) finds that his negative selection results differ from the positive selection of Chiquiar and Hanson (2005) only due to the undercounting of the low-skilled migrants in the data used by the latter.

Self-selection has implications for sending and receiving economies. For sending countries the literature on brain drain argues that the dominant channel is through the deprivation of the sending countries of skills required locally, thereby subtracting from their growth potential (Bhagwati and Hamada, 1974). In contrast, the brain gain literature finds that due to a positive emigration probability, individuals tend to obtain more human capital. Since only a small fraction of the population eventually emigrates, the sending country has a higher supply of human capital than under trivial emigration probability (see Commander et al., 2003, Batista et al., 2012). Gibson and McKenzie (2012) suggest that high probability of migration affects the choice of education field more than the level of education.

For receiving economies self-selection is linked to labor market outcomes: earnings assimilation (Adsera and Chiswick, 2007 for migrants in the EU; Berman et al., 2003 for migrants in Israel; Friedberg, 2000 and Moraga, 2011) and occupation attainments (Mattoo et al., 2008 for the US; Kostenko et al., 2012 for Australia; and Turner, 2010 for Ireland).

The bulk of the existing research on earnings assimilation suggests that upon arrival migrants face a significant wage gap compared to locals in the same occupation with similar observables. Adsera and Chiswick (2007) find the gap to be 40% on average, though it widens for migrants born outside the EU and varies across destination countries. Berman et al. (2003) find that wages for migrants from the former Soviet Union in Israel converge for those in the upper part of the occupation distribution, and the rate of convergence is closely linked to the knowledge of Hebrew. They report no convergence in wages, irrespective of Hebrew proficiency, for occupations at the bottom of the occupation distribution.

Mattoo et al. (2008) find that in the US labor market, migrants from Latin America and Eastern Europe are more likely to end up with low-skilled jobs than migrants from Asia and developed countries with similar characteristics. The authors explain this variation by low or poorly transferable skills obtained in certain source countries, as well as the selective US immigration policy. Poor quality or low transferability of human capital is also related to expenditures on tertiary education and the use of English as a medium of instruction in source countries. The US immigration policy matters as migrants from certain countries are more likely to be admitted through family reunification and the visa lottery programmes, whereas migrants from other countries use the labor market channel.

At all stages of emigration and employment, the literature acknowledges the presence of "network effects" when family members and friends share their networks, thus reducing emigration and search cost. Jackson (2008) develops game theoretical tools for the study of social networks and Munshi (2003), McKenzie and Rapoport (2010), Ioannides and Loury (2004) provide relevant empirical evidence.

3.3 Survey design

The current study is based on a tailor-made survey. The population of interest was defined as persons in the labor force, males and females, aged 15–59, who reside in non-institutionalized dwellings in settlements with a population of 50,000 and more. The decision to ignore rural population was dictated by several considerations, including difficulties in achieving adequate coverage as well as the fact that Ukrainian domestic migration from rural to urban areas has historically been significant. As a consequence, earlier evidence suggests that the bulk of potential external migrants are resident in urban areas.[2]

Due to the differences in cultural and historical backgrounds across various geographical parts of Ukraine as well as infrastructural diversities across settlements of various sizes, we stratified the sample by region and town size. Four geographical regions (West, Center and North, East, and South) and four town sizes (50–100K, 100–200K, 200–500K, 500K–1mln residents) resulted in a 4-by-4 stratification map + large cities (> 1mln residents) as a separate stratum. The population weights on gender and 10-year age brackets were calculated using data from the Ukrainian State Statistical Office.

The data were collected by means of direct interviews with households between August and October 2011. Specific search routes were selected to maximize the distance between each pair of sampling points in towns that had been randomly chosen within a particular stratification cell. Depending on availability within a selected household, responses were collected from one randomly chosen member without any (external) migration experience and from all members with such experience. To be considered a person with migration-related experience, a person had to meet one of the following criteria at the time of the interview:

1. Be residing, working or studying abroad (currently abroad category, CA);

2. Have been abroad for the purpose of residence, employment or education in the most recent three years (return migrant category, RM);

[2]There is one obvious exception. In the western part of Ukraine, much of the external migration to neighboring countries has been from rural areas.

3. Be planning to go abroad for the purpose of residence, employment or education in the next 12 months (prospective migrant category, PM).

When it was not possible to gather information about a person directly, we had to rely on other household members most knowledgeable of the matter.[3] Altogether, we obtained information on 6676 individuals from 5985 households living in 63 towns of Ukraine.

3.4 Description of collected data

3.4.1 Descriptive statistics

Table 1 provides some basic descriptive statistics on the sample. The unweighted (weighted in brackets) distribution of migrants across the categories is as follows: currently abroad 409 (369), returnees 216 (266), prospective migrants 320 (320), and non-migrants 5739 (5720). In what follows we weight observations to generalize results for the whole urban population. We classify respondents from the categories currently abroad and returnees as migrants. We omit the prospective migrants' category from most of the discussion because by definition individuals in this category do not belong to either migrants or non-migrants.

Between the four migration categories in Table 1 we test for six types of selection: currently abroad vs returnees, currently abroad vs prospective, currently abroad vs non-migrants; returnees vs prospective, returnees vs non-migrants; and prospective vs non-migrants. Within each migration category we test for selection on gender. The results of these tests are provided in Tables C.1.2 and C.1.3.

CA, RM and PM categories are heavily dominated by males compared to the non-migrants. PM is the youngest category with an average age of around 31 years. This compares to about 37 years in the remaining three categories. Despite the same mean age, the age structure differs significantly across the categories. Quite surprisingly, there is little gender difference in mean age and age structure within each category, as is illustrated in Table C.1.2.

[3]We understand that this might introduce imprecision into the information collected, although this is a better solution than entirely disregarding these individuals.

Table 1: Basic descriptive statistics.

Variable	Currently abroad N	%	Returnees N	%	Prospective N	%	Non-migrants N	%
Male	245	66.4	196	73.7	199	62.2	2534	44.3
Age								
mean	36.8		37.3		31.0		37.5	
s.d.	11.4		10.3		11.7		12.8	
15-19	9	2.4	4	1.5	56	17.5	536	9.4
20-29	121	32.8	76	28.6	121	37.8	1342	23.5
30-39	82	22.2	66	24.8	56	17.5	1229	21.5
40-49	90	24.4	85	32.0	58	18.1	1286	22.5
50-59	67	18.2	35	13.2	28	8.8	1328	23.2
Marital status								
Single	97	26.3	46	17.3	151	47.2	1435	25.1
Married	235	63.7	173	65.0	124	38.8	3270	57.2
Cohabitation	14	3.8	14	5.3	14	4.4	248	4.3
Divorced	20	5.4	29	10.9	27	8.4	563	9.8
Widowed	3	0.8	4	1.5	4	1.3	204	3.6
Education								
Primary	0	0.0	0	0.0	0	0	12	0.2
Basic secondary	4	1.1	1	0.4	11	3.4	230	4.0
Complete secondary	33	8.9	16	6.0	50	15.6	790	13.8
Vocational	82	22.2	114	42.9	60	18.8	1435	25.1
Basic higher	86	23.3	65	24.4	79	24.7	1441	25.2
Complete higher	158	42.8	67	25.2	119	37.2	1786	31.2
Candidate of sciences	5	1.4	3	1.1	1	0.3	25	0.4
Doctor of sciences	1	0.3	0	0.0	0	0	1	0.0
Total weighted	369		266		320		5720	

There is also significant variation in the marital status and education by groups and gender. More than 40% of the currently abroad migrants have completed higher education (Master's degree or equivalent) compared to 31% amongst the non-migrants, thus suggesting positive selection by education level. The returnees are more likely to be married males from the middle part of the education distribution: the share with vocational training is 42.9% relative to 22.2% and 25.1% in the currently abroad and non-migrant categories respectively. This suggests medium-level selection for the returnees.[4] Respondents with higher education, except return migrants, are more likely to be women.

The emigration rate is 10% with an associated 95% confidence interval $[9.28, 10.72]$. Table C.1.1 and Figure C.2.1 illustrate large variation in emigration rates and destination countries across administrative regions. The Western part of Ukraine plus Odesa and Lugansk regions are heavily affected by emigration. However, no particular spatial pattern is observed for the rest of the country. The lowest migration rate is 1.8% in Zhytomyr region and the highest is 41.9% in Ternopil region. There is a clear trend for the respondents from the Western regions to go to the EU countries, and for the Eastern part to choose the former Soviet Republics, primarily Russia. The highest emigration rate to the EU is 75.2% in Lviv region. Lugansk region has the lowest emigration to the EU, but the highest emigration to Russia (89.7%). Vinnytsia and Odesa regions have large fractions of migrants, 60.3% and 45% respectively, traveling to other countries (mainly the US, Canada and Israel).

The most frequently chosen destination is Russia with 40.1% and 50.2% of the currently abroad and return migrants respectively (see Table 2). Among the EU destinations Italy, Poland and Germany are the three most frequent while the USA, Israel and the UAE are the three most favored destinations in the rest of the world category.

The existing research unambiguously suggests that the primary reason for migration is the difference in wage rates net of migration costs. In our survey 76.7% of all the migrants chose options "better pay" and "better employment opportunities" as their primary reason for emigration. Table 3 summarizes data on the self-reported average monthly income.[5]

[4] Medium-level selection into return migration is also found in a number of other papers; see Martin and Radu (2012) for a review of relevant empirical studies.

[5] For the income (remittance) question we obtain relatively high response rates, namely 58.3% (84.8%) for the currently broad, 77.4% (94.3%) for the returnees, 76.3% for the prospective migrants and 86.9% for the non-migrants.

Table 2: Chosen destinations.

	Currently abroad		Returnees	
Destination	N	%	N	%
Russia	144	39	134	50.2
Italy	36	9.8	20	7.6
Poland	25	6.9	31	11.7
Germany	23	6.2	15	5.8
USA	20	5.3	5	1.9
Israel	15	4	1	0.3
Spain	12	3.3	7	2.7
Czech Republic	9	2.4	12	4.5
Greece	7	1.8	8	2.9
Portugal	7	1.8	2	0.6
UAE	5	1.5	1	0.5
UK	5	1.3	3	1.1
Other	62	17	27	10.1
Total	369		266	

An average non-migrant working in the manual labor sector in Ukraine reports earnings of USD 554.9, while a migrant working in the same sector abroad reports more than twice that amount: USD 1294.8 in the EU15 and USD 2042.6 in the EU10. The income gap persists and increases further up the occupation ladder.[6] The highest income sector in Ukraine pays less than the least rewarding sector abroad. This implies that for a skilled Ukrainian with relatively low migration costs it may be attractive to take up an unskilled job abroad, possibly avoiding occupation search and integration costs.

Table 3: Self-reported average monthly income for 6 months prior to survey date, PPP-adjusted USD.

Sector	EU15	EU10	Russia	ROW	Ukraine
Manual	1294.8	2042.6	1717.8	1871.4	554.9
Specialized manual	1231.6	1689.9	1796.4	2048.2	672.4
Highly-skilled	1830.1	1563.3	2092.4	2097.5	737.9
Narrow highly-skilled	4538.6	2551.0	1489.1	3314.1	876.8
Administrative	1924.1	2506.1	2701.1	2677.6	1158.0

The data show that return migrants do not necessarily get a "migration premium" for their experience abroad, a phenomenon also found by other authors (Co et al., 2000 for

[6] The discussion must not omit the existence of the significant unofficial sector in Ukraine, which the current survey design did not aim to measure.

returning Hungarians; and Ambrosini and Peri, 2012 for returning Mexican migrants).

As regards the reasons for returning home, 55% of the return migrants chose "personal reasons" and 31% reported reasons relating to "employment contract expiration" and "conclusion of their education." It thus seems that the decision to return is not entirely at the migrants' discretion, a fact that potentially reduces the selection bias into return migration. This finding is similar to Gibson and McKenzie (2012), who find that migrants in their survey return to their countries of origin mainly for family or lifestyle reasons.

Around 40% of the migrants send or have sent remittances home, the rate being slightly higher for the currently abroad than for the returnee category (45% and 40% respectively). The average weighted amount remitted is USD 535 for the currently abroad and USD 760 for the returnee groups.[7]

The survey instrument also enables the investigation of whether human capital improving actions have been taken by respondents as a way of raising their probability of emigration. This is, of course, a central proposition in the wider brain gain literature (see, *inter alia*, Beine et al., 2008 and Beine et al., 2011). Respondents were asked if they had taken additional education or training to improve their chances of emigration and their answers are summarized in Table 4.[8]

Table 4: Answers to question "Have you tried to improve your chances of emigration by any of the following?"

	Currently abroad		Returnees		Prospective		Non-migrants	
	N	%	N	%	N	%	N	%
Additional years of schooling	11	3	0	0	7	2.2	16	1.6
Language classes	59	16	14	5.3	61	19.1	115	11.3
Professional skills building	55	15	20	7.5	33	10.3	48	4.7
Private classes	14	3.8	0	0	4	1.3	16	1.6
Preparation for SAT	16	4.3	6	2.3	7	2.2	10	1
Have not tried	200	54.2	216	81.2	176	55	730	71.9
Other	3	0.8	0	0	9	2.8	26	2.6
Do not know	49	13.3	10	3.8	23	7.2	54	5.3
Respondents in category	369		266		320		1015	

The brain gain effect is present if among prospective and non-migrant categories edu-

[7]The remittances were predominantly in cash.

[8]In the non-migrant category this question was asked to those who ever considered emigrating. This reduces the number of answers to 1015.

cation or skill-improving choices are selected. Table 4 shows that there is some evidence of active acquisition of additional years of schooling. Prospective migrants are more active in acquiring human capital than non-migrants. In particular, 19% of them took language classes, 10% had improved their professional skill and 2% took additional years of schooling compared to 11.3%, 4.7% and 1.6% amongst non-migrants respectively. However, this is relatively weak evidence for the presence of any brain gain channel.

To explore the conjecture further, we looked at whether emigration rates in particular localities were associated with emigration improving actions. To that end, we estimated a simple regression relating whether an individual had tried to improve her chances to emigrate measured by any option from Table 4 to the city (town) level emigration rate where that individual lived, as well as the region (oblast) level emigration rate. We also included a vector of individual characteristics: age, education, experience, marital status, etc. We found no significant association between the acknowledged acquisition of additional human capital and emigration rates on the city and/or region levels.

3.4.2 Labor market outcomes: defining downshifters

A striking feature of the labor market performance of the Ukrainian migrants is that they commonly downshift and take up work for which they are seemingly overqualified. A standard definition of overeducation relates educational attainment at home to a labor market match abroad (Mattoo et al., 2008). The former is considered to reflect unobservable skills while setting an aspiration level. To gauge the nature of the match, our survey instrument contains a five-point ranking of the skill intensity of respondents' current occupations. These categories are manual labor, specialized manual labor, general highly-skilled, specialized highly-skilled, and administrative. A downshifter is defined as someone for whom one of the following holds:

1. Involuntarily unemployed;

2. Employed in manual labor if the skill level is medium or high;

3. Employed in specialized manual or manual labor if the skill level is high.

In Table 5 we tabulate the distribution of occupations by educational attainments. For the latter, the broad categories are applied. Low-skilled people are defined as having primary and/or basic secondary education; medium-skilled respondents with complete secondary and vocational education, and highly-skilled individuals with tertiary education. The shaded grey areas in Table 5 highlight the incidence of overeducation. It appears that 43% of medium-skilled and 56% of highly-skilled respondents downshifted abroad. In total 288 individuals or just over 44% of the migrants downshifted.

Table 5: Occupational distribution of respondents.

	Low-skilled			Medium-skilled			Highly-skilled		
	Migrants		Non-	Migrants		Non-	Migrants		Non-
	Before	After	migrants	Before	After	migrants	Before	After	migrants
Unemployed	.	.	8	38	3	193	28	4	162
Manual labor	2	1	24	48	71	329	19	79	132
Spec. manual labor	.	.	26	102	144	972	84	131	663
Gen. highly-skilled	.	.	3	9	7	224	110	86	1318
Spec. highly-skilled	1	.	.	.	.	10	24	25	141
Administrative	.	2	1	5	5	35	10	13	282
Study	2	2	163	12	8	224	73	22	229
Other	.	.	17	31	7	239	37	25	325
Downshifters	.	.	8	86	74	522	131	214	957
Non-downshifters	5	5	234	159	171	1704	254	171	2295

Notes: The shaded areas show baseline downshifters. "Low" denotes primary and basic secondary education, "Medium" stands for complete secondary and vocational, and "High" is for tertiary. "Before" refers to the period prior to emigration (employment in Ukraine), "After" refers to after emigration (employment abroad). A dot means zero value.

Although the baseline definition is widely used in the literature, we have reasons to believe that its core assumption, namely that education is a good signal of an individual's labor market skills, may be questioned in the context of many transition economies, such as Ukraine. Aside from the fact that education may not proxy unobservable skills well (see for example Heckman and Rubinstein, 2000) there is also evidence that in transition countries, the inherited system of education has not been well adapted to the needs of a market economy. This implies that the signal from education to skills has become less robust than might normally be the case, suggesting that the conventional measure may potentially be misleading.

A novel feature of our survey is that we collected information on migrants' labor market statuses before and after emigration. This allows us to relate migrants' occupational choices abroad to their prior occupations in Ukraine. We thus create two dummy variables; *shift_abr* and *shift_ua* for whether a person is a baseline downshifter abroad and in Ukraine respectively. We tabulate these two dummies in Table 6. The tetrachoric correlation coefficient between them is 0.69 with standard error 0.04, $\rho_T = 0.69(0.04)$.

Table 6: Tabulation of overeducation abroad vs *overeducation in Ukraine,* $\rho_T = 0.69(0.04)$.

		shift_abr		
		No	Yes	Total
shift_ua	No	299	119	418
	Yes	48	169	217
	Total	347	288	635

Slightly less than 60% of downshifters abroad (169 out of 288) had not been well matched in Ukraine prior to emigration. This implies that downshifting abroad may not be understood simply in terms of migrants' inability to find appropriate work or other related explanations. With this in mind, we provide an alternative definition of downshifting, which compares employment abroad to employment in Ukraine. A migrant is defined as an alternative downshifter (dummy variable *shift*) if one of the following two criteria is satisfied:

1. Involuntarily unemployed abroad if employed in Ukraine;

2. Employed in an occupation abroad that is below the pre-migration level.

When applying this filter, the number of downshifters narrows from 288 to 116, or less than 20% of 635 migrants in our dataset. This suggests that the baseline estimate with its underlying and strong assumptions concerning the relationship between education and skills may be inappropriate in the context of this study.

Table 7: Downshifters: the baseline vs *alternative definitions,* $\rho_T = 0.8(0.04)$.

		shift_abr		
		No	Yes	Total
shift	No	342	177	519
	Yes	5	111	116
	Total	347	288	635

The tabulation in Table 7 shows that the baseline definition is more likely to consider someone a downshifter when the alternative does not than the other way round.[9] This supports the conjecture that the baseline definition of a downshifter can sometimes be overly optimistic.

Table 8: Downshifters: alternative vs *downshifter in Ukraine,* $\rho_T = -0.24(0.07)$.

		shift_ua		
		No	Yes	Total
shift	No	327	193	519
	Yes	91	25	116
	Total	418	217	635

As can be seen in Table 8, 22% of non-downshifters in Ukraine end up as alternative downshifters abroad, whereas only 11% of downshifters in Ukraine are alternative downshifters abroad. Baseline shifters in Ukraine are less likely to be alternative shifters abroad.

3.4.3 Network effects

The survey design enables us to test for the presence of three types of family network effects:

1. Current migrants affect migration decisions of non-migrants (variable $migr\_fam$).
2. Destinations of new migrants are affected by destinations of old migrants from the same family (variable $dest\_fam$).
3. Occupation choices of new migrants are affected by occupations of old migrants from the same family (variables $shift\_fam\_abr$ and $shift\_fam$).

Network effects have been found in many contexts to be important in explaining choice of destination and, in some cases, choice of occupation. For example, a family with a migrant possesses more migration-related information and that may mean that its members are more likely to emigrate. Migrants from the same family are more likely to

[9] Quite expectedly, out of those 177 people, the majority (over 80%) had a poor labor market match already in Ukraine.

choose similar destinations and occupations because sharing a network of contact reduces search costs. Dummy variable $migr\_fam$ is created for whether a family already has a migrant.

There are 57 households with at least 2 migrants. Migrants from 45 of these households traveled to the same country. We create a dummy variable $dest\_fam$ for migrants in the same destination. We wish to test if the presence of two migrants from the same family in the same destination has any effect on their occupation choices. The simple tabulation of variables $shift\_abr$ and $dest\_fam$ in Table 9 does not give a clear answer to this question.

Table 9: Tabulation of migrants' destinations and occupation choices. The number of households is in brackets.

		dest_fam		
		No	Yes	Total
shift_abr	No	9 (5)	53 (26)	62 (31)
	Yes	8 (7)	34 (19)	42 (26)
	Total	17 (12)	87 (45)	104 (57)

Finally, we create dummy variables $shift\_fam\_abr$ and ($shift\_abr$) for whether a migrant has another migrant in the family, who is a baseline (alternative) downshifter. Table 10 shows that if one out of two migrants is an alternative downshifter, the other is certain to be an alternative downshifter. Similar intuition holds for classical downshifters.

Table 10: The tabulation of $shift\_abr$ against $shift\_fam\_abr$ and ($shift\_fam$).

		shift_fam_abr (shift_fam)		
		No	Yes	Total
shift_abr	No	81 (61)	11 (16)	92 (77)
	Yes	0 (0)	12 (27)	12 (27)
	Total	81 (61)	23 (43)	104 (104)

We now proceed to formalize the model with unobserved individual heterogeneity, estimate it and compare the estimates between the baseline and alternative definitions of downshifters.

3.5 Econometric model and estimation

3.5.1 The model and identification

The underlying theoretical framework is based on random utility foundations. We assume that each individual has two distinct but inter-related decisions to make – to emigrate from Ukraine and to downshift abroad taken in the sequential order as shown in Figure 1.

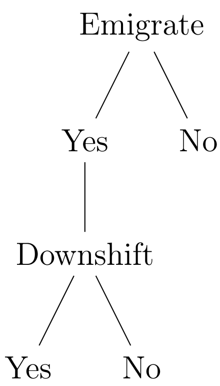


Figure 1: Decision tree of an average individual in the sample.

The decision to emigrate is determined by a selection equation and the decision to downshift is modeled with an outcome equation. We formalize each equation in turn.

Let w_i^{ua} be the wage of individual i in Ukraine and w_i^{abr} be the wage abroad. In order to emigrate, a worker must cover emigration costs c_i, which are related to the search for an employer, learning a foreign language, applying for a visa, and buying a plane ticket. The wage and cost functions are assumed to be of additive form:

$$w_i^{ua} = x_i\alpha_1 + x_i^{ua}\alpha_2 + \mu_i + \epsilon_i^{ua}, \tag{3.1}$$

$$w_i^{abr} = x_i\beta_1 + x_i^{ua}\beta_2 + \mu_i^{abr} + \mu_i + \epsilon_i^{abr}, \tag{3.2}$$

$$c_i = x_i\theta_1 + x_i^{ua}\theta_2 + \mu_i^{abr} + \epsilon_i, \tag{3.3}$$

where x_i is a row vector of individual specific characteristics, which remain largely unchanged irrespective of a person's location: gender, age, education, etc. x_i^{ua} is a row vector of individual characteristics, which are specific to Ukraine, such as the region of origin and the city of residence. μ_i^{abr} captures destination country heterogeneity: immigration policy, certification requirements to foreigners, difficulty of finding a job, technology level, etc. μ_i is unobserved individual heterogeneity, which captures unobserved skills and motivation; ϵ_i^{ua}, ϵ_i^{abr} and ϵ_i are error terms from respective equations.

Putting the issues of non-reporting aside, wages of Ukrainians in Ukraine are well observed. However, the wages of Ukrainian migrants abroad are not observed for non-migrants. For this reason w_i^{abr} is not a function of characteristics at destination.[10]

Given wages and migration costs, the individual maximization problem is:[11]

$$\max \{w_i^{abr} - c_i, w_i^{ua}\} \tag{3.4}$$
$$\text{s.t.: equations (3.1), (3.2) and (3.3).}$$

In other words, a person emigrates ($y_{1i} = 1$) from Ukraine if $w_i^{abr} - w_i^{ua} - c_i > 0$ and stays ($y_{1i} = 0$) otherwise. The probability that a random individual emigrates is:

$$Prob(\text{emigrate}) = 1 - F\left[-\left(x_i(\beta_1 - \alpha_1 - \theta_1) + x_i^{ua}(\beta_2 - \alpha_2 - \theta_2)\right)\right], \tag{3.5}$$

where $F(\cdot)$ is a distribution function of the composite error term $\left(\epsilon_i^{abr} - \epsilon_i^{ua} - \epsilon_i\right)$. y_{1i} is a binary variable, which equals 0 if i's decision is to stay and 1 if i's decision is to emigrate.

For migrants the skill intensity of occupations can be described by the following equations:

$$\begin{aligned} y_i^{ua} &= x_i\gamma_1 + x_i^{ua}\gamma_2 + x_i^{abr}\gamma_3 + \mu_i + \nu_i^{ua} \\ y_i^{abr} &= x_i\pi_1 + x_i^{ua}\pi_2 + x_i^{abr}\pi_3 + \mu_i + \nu_i^{abr}, \end{aligned}$$

where y_i^{ua} and y_i^{abr} are continuous variables, which define the skill intensity of occupations in Ukraine and abroad, respectively. y_i^{ua} and y_i^{abr} are functions which assign non-negative real values to each occupation: more skill intensive occupations get higher values. x_i^{abr} is a row vector of individual characteristics abroad, such as the duration of stay, visa status,

[10] Similar wage functions are mentioned by Greene (2012) on p. 685 and p. 879.

[11] Optimization problem (3.4) assumes that a worker's utility function is linear in income, implying that a certain level of wage gap is equally attractive to high and low earning workers in Ukraine. An alternative approach is to assume a log-linear utility function, whereby the same level of wage gap becomes less attractive as a worker's wage in Ukraine increases. Grogger and Hanson (2011) elaborate on the differences between these two approaches.

the nationality of employer and destination effects.

$(y_i^{ua} - y_i^{abr})$ measures the extent of alternative downshifting. Conditional on country specific covariates, it shows how the skill intensity of an occupation in Ukraine is different from the skill intensity of an occupation abroad for the same i. An individual is defined an alternative downshifter if $(y_i^{ua} - y_i^{abr}) > 0$ and a non-downshifter otherwise. Let y_{2i} be a dichotomous variable, which equals 0 if i is not a downshifter and 1 if i is a downshifter.

The probability of being a downshifter is:

$$Prob(\text{downshift}) = 1 - G\left[-\left(x_i(\gamma_1 - \pi_1) + x_i^{ua}(\gamma_2 - \pi_2) + x_i^{abr}(\gamma_3 - \pi_3)\right)\right], \tag{3.6}$$

where $G(\cdot)$ is a distribution function of the composite error term $\left(\nu_i^{ua} - \nu_i^{abr}\right)$.

An appealing feature of identities (3.5) and (3.6) is that the unobserved heterogeneity component, μ_i, is no longer there because it has been differenced out. Until this point no particular distribution assumption has been imposed on the error term structure. Let $\epsilon_{1i} = \left(\epsilon_i^{abr} - \epsilon_i^{ua} - \epsilon_i\right)$ and $\epsilon_{2i} = \left(\nu_i^{ua} - \nu_i^{abr}\right)$ have a bivariate normal distribution with zero means and correlation coefficient ρ, $(\epsilon_{1i}, \epsilon_{2i}) \sim N(0, 0, \sigma_1^2, \sigma_2^2, \rho)$. Under this assumption the model is known to be probit with sample selection (see Van de Ven and Pragg, 1981). Using the above probabilities and individual weights, ζ_i, the log-likelihood function for the same of size N is:

$$\begin{aligned} \ln L &= \sum_{i=1}^{N} \zeta_i \cdot \Big\{ (1 - y_{1i}) \cdot \ln \Phi\left(-(x_i\alpha + x_i^{ua}\beta)\right) + \\ &+ y_{i1}(1 - y_{2i}) \cdot \ln\left[\Phi\left(x_i\alpha + x_i^{ua}\beta\right) - \Phi_2(x_i\alpha + x_i^{ua}\beta, x_i\gamma + x_i^{ua}\theta + x_i^{abr}\pi, \rho)\right] \\ &+ y_{1i}y_{2i} \cdot \ln \Phi_2(x_i\alpha + x_i^{ua}\beta, x_i\gamma + x_i^{ua}\theta + x_i^{abr}\pi, \rho) \Big\}, \end{aligned} \tag{3.7}$$

where $\alpha = \beta_1 - \alpha_1 - \theta_1$, $\beta = \beta_2 - \alpha_2 - \theta_2$, $\gamma = \gamma_1 - \pi_1$, $\theta = \gamma_2 - \pi_2$, $\pi = \gamma_3 - \pi_3$ are identified parameters.

Equation (3.7) requires that the variables in the outcome equation are a subset of variables from the selection equation. This is required to avoid perfect multicollinearity between the variables in the outcome equation and the selection term carried over from the selection equation. Our exclusion restriction consists of variables *unempl_y_ua*, *unempl_o_ua* and *migr_fam*. The migrant's unemployment status is expected to affect the probability of emigration, but does not directly affect the labor market outcome outside Ukraine. Variable $migr\_fam$ increases the probability of emigration via network effects, but does not affect downshifting abroad because we create variables $dest\_fam$ and $shift\_fam$ which proxy for employment related network effects. The exact definitions of the covariates included in each vector are given in Table C.1.6.

Some variables are defined for all respondents and can thus be included in both equations. However, many variables are defined for migrants only. This includes the choice of a destination country, employment details abroad, visa status abroad, and migration aid received. These variables are included in the outcome equation but excluded from the selection equation because they are not defined for non-migrants.

3.5.2 Estimation results

We estimate the log-likelihood function (3.7) for the baseline and alternative definitions of downshifters. We thus have the estimates from two selection equations and two outcome equations. The estimates of the marginal effects are illustrated in Table C.1.4. Since the definition of a migrant is the same, the estimates of the selection equations are almost identical. We first go over the estimates of the marginal effects of the selection equation, and then discuss the estimates of the outcome equation. Table C.1.4 is constructed to give a clear illustration what variables are included in both equations and what variables serve as an exclusion restriction.

An average migrant is a married male most likely from the West or South of Ukraine, who typically comes from the middle (vocational training) or upper (Bachelor's degree and above) parts of the education distribution. There is a concave relationship between the probability of emigration and age; the emigration probability increases until the age of 35.5 and then slowly declines.

The household income adds to the emigration probability in a non-linear manner. Being from an average and above average income family adds 5% and 12.7% to the emigration probability respectively compared to the below average category. This suggests that emigration may be more of an investment in a better future than just an escape from poverty. This finding is consistent with other empirical studies on poverty traps and the effect of income on emigration probability (Mayda, 2010; Pedersen et al., 2008).

A discussion on the endogeneity of income is relevant here. We do not use the level of income but widely defined income categories, namely below average, average, and above average. It is likely that a person saved up or even borrowed financial resources prior to emigration. This would increase the level of income but not change the income category. We therefore think that it is reasonable to assume that at the time of emigration the migrant's family income category is given.

City size, as a proxy for average income, reveals a concave shape on the probability of emigration. Respondents from small and large cities are less likely to be migrants compared to respondents from medium-sized cities. This effect works through the budget constraint, whereby in small towns people have less economic opportunities and thus less resources to cover the costs of emigration. In contrast, better economic opportunities in large cities deter residents from emigration.

It is surprising that after some threshold additional years of education have no effect on the probability of emigration. Moving from category $educ2$ to $educ3$ does not increase (in a statistical sense) the chances to emigrate. This suggests two complementary conjectures. First, labor demand abroad may be skewed to particular types of skills and occupations. Second, more education does not necessarily translate to skills useful outside Ukraine.

There is no difference in the probability of emigration between the unemployed respondents below and above the age of 30. Having a migrant in the family increases the probability of emigration 11.4% compared to a family without a migrant.

The estimates of the determinants of downshifting differ between the baseline and alternative definitions. The comparison narrows down to how deficiencies in the Ukrainian education system affect labor market outcomes. The major difference is in the signs of the variable $shift\_ua$. A migrant who was a baseline downshifter in Ukraine is 14% more likely to downshift abroad. It may be explained by the fact that individual unobserved heterogeneity persists over time and across destinations.

In contrast, a baseline downshifter in Ukraine is 2.5% less likely to be an alternative downshifter abroad. It is so because when a respondent is matched in Ukraine her unobserved skills are partly revealed. Once she emigrates, the match persists. If a non-downshifter in Ukraine becomes a downshifter abroad, then this points to poor cross border transferability of human capital obtained in Ukraine.

A typical downshifter, baseline or alternative, is a single male aged below 40, not a head of a household and not coming from a large city. The probability of downshifting increases with the level of education and decreases with the knowledge of a local language or English. Such factors as family income, residence status or reason for emigration are not significant determinants of alternative downshifting. The probability of alternative downshifting significantly increases with the presence of another downshifter abroad.

Ukrainians are equally likely to be baseline downshifters irrespective of destination, whereas alternative downshifters are more likely to be migrants in the countries of the former Soviet Union and the European Union as compared to the baseline category. The estimates indicate that accounting for other covariates the probability of downshifting is statistically the same for migrants to the countries of the former Soviet Union and the European Union. The estimates refer to probabilities but not to the intensity of downshifting. For example, a teacher of the Ukrainian literature working as a cleaner in Italy is a more severe downshifter than the same teacher working as a shop clerk in the suburbs of Moscow. However, the probabilities of being a downshifter (irrespective of the intensity) might be similar.

3.5.3 Robustness check

We construct the robustness check of the estimates in three steps. We first validate the bivariate normal distribution (goodness-of-fit test), then we endogenously determine threshold $\bar{p}$, which maximizes the number of correctly classified observations. Finally, we apply the semi-nonparametric estimator of Luca (2008) and Luca and Perotti (2011) to see how the estimates vary in case the distribution is misspecified.

Using the Doornik-Hansen test, we do not reject the null hypothesis that the error terms $\hat{\epsilon}_{1i}$ and $\hat{\epsilon}_{2i}$ have bivariate normal distribution.[12] This validates the main underlying assumption behind the parametric estimation and the maximum likelihood estimator is thus consistent and asymptotically efficient (Luca, 2008; Martins, 2001; Luca and Perotti, 2011). As a way to visualize the result, we draw two quantile-quantile plots in Figure

[12] H_0: $\hat{\epsilon}_{1i}$ and $\hat{\epsilon}_{2i}$ have bivariate normal distribution; H_A: H_0 is not true. The test statistics is $\chi^2_4 = 4.330$, Prob $> \chi^2_4 = 0.3631$.

C.2.2. In Figure C.2.2(a) we plot the estimated error term from the selection equation against the estimated error term from the outcome equation. In Figure C.2.2(b) we show a similar quantile-quantile plot for the theoretical bivariate normal distribution with a respective mean and variance-covariance structure. Since the two plots are similar, the error terms are likely to come from the bivariate normal distribution.

Given the validity of the estimates in Table C.1.4, we count the percentage of correctly classified observations. In view of the fact that the share of successes in the data is only 10% it might be unreasonable to exogenously choose 0.5 as the cutoff probability. Instead, we choose such $\bar{p}$, which minimizes the number of misclassified cases (false positives and false negatives):

$$\max_{\bar{p}} \; S(\bar{p}) = -\sum_i |y_i - \hat{y}_i(\bar{p})| \,, \tag{3.8}$$

where y_i is a binary variable from the data; $\hat{y}_i(\bar{p})$ is a binary variable, which equals 1 if $\hat{p} > \bar{p}$ and 0 otherwise. The function $S(\bar{p})$ for the selection and outcome equations (alternative definition) is depicted in Figure 2. Hence, using thresholds $\bar{p}_m = 0.45$ for the decision to emigrate and $\bar{p}_d = 0.48$ for the decision to downshift we have slightly more than 10% of misclassified cases.

Table C.1.5 contains the semi-nonparametric estimates of equation (3.7). The majority of estimates do not change their signs or marginal effects, though the significance of the estimates changes quite significantly. This happens because the semi-nonparametric estimator is less efficient when the distributional assumption is correctly specified, leading to inflated standard errors.

3.6 Conclusion

Our paper focuses on the patterns of self-selection and labor market outcomes among Ukrainian migrants. It confirms significant selection on gender, education, pre-migration income, region, and city of origin in Ukraine. However, when it comes to labor market outcomes of migrants, we find clear patterns of overeducation and occupational downshift. Indeed, over 45% of the migrants in the collected sample have a level of education that

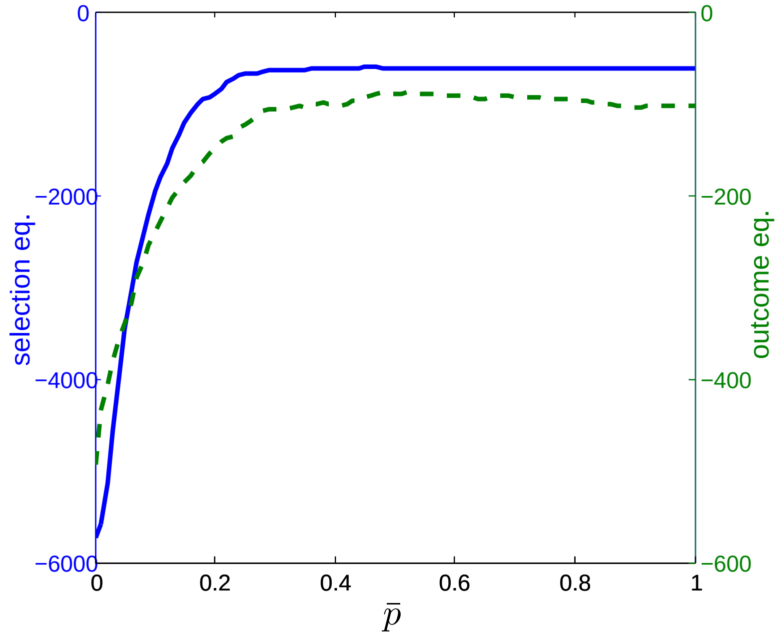


Figure 2: Visualization of function $S(\bar{p})$ from maximization problem (3.8).

by far exceeds their job requirements. This might be attributed to the migrants having a disadvantageous position in local labor markets, whether through discrimination or informational shortcomings, or it might be related to their "true" attributes which are not necessarily well captured by their education level. To address this, we look at labor market outcomes before and after emigration. We find that a person who had previously downshifted in Ukraine is about 14% more likely to be a downshifter abroad. This suggests that education is indeed a noisy signal of individual unobserved ability and an alternative measure may be required. We set up a simple model that focuses on the migration and downshift decisions sequentially, accounts for individual unobserved heterogeneity and estimate it using our survey data.

The title of our paper asks whether Ukrainian migration has been mainly about skilled (brain) or unskilled (brawn) migration. Our answer is qualified. For those migrants currently abroad, the picture is one where in terms of educational attainments a clear majority has some level of tertiary education. The profile appears biased towards skills. However, when looking at what Ukrainian migrants do when they emigrate, a significant share of them work in occupations that match poorly to their prior educational attainments. This suggests that migration involves downshifting. But this picture is itself somewhat misleading as our analysis shows that a significant number of these downshifters had already downshifted at home prior to emigration. We consider that drawing strong conclusions about the efficiency of occupation-education matching may not be warranted. Rather, what may be a more promising avenue of enquiry, particularly from a policy perspective, is to consider why the educational attainments of Ukrainians have such a weak link to labor market outcomes. The answer is likely to lie in the deficiencies of the current educational system and the limited adaptation that has been made to the needs of the labor market in a market economy whether at home or abroad.

Bibliography

Adsera, A., Chiswick, B., 2007. Are there gender and country of origin differences in immigrant labor market outcomes across European destinations? Journal of Population Economics 20 (3), 495–526.

Ambrosini, J., Peri, G., 2012. The determinants and the selection of Mexico–US migrants. The World Economy 35 (2), 111–151.

Batista, C., Lacuesta, A., Vicente, P., 2012. Testing the 'brain gain' hypothesis: micro evidence from Cape Verde. Journal of Development Economics 97 (1), 32–45.

Beine, M., Docquier, F., Özden, Ç., 2011. Diasporas. Journal of Development Economics 95 (1), 30–41.

Beine, M., Docquier, F., Rapoport, H., 2008. Brain drain and human capital formation in developing countries: Winners and losers. The Economic Journal 118 (528), 631–652.

Berman, E., Lang, K., Siniver, E., 2003. Language-skill complementarity: returns to immigrant language acquisition. Labour Economics 10 (3), 265–290.

Bhagwati, J., Hamada, K., 1974. The brain drain, international integration of markets for professionals and unemployment: a theoretical analysis. Journal of Development Economics 1 (1), 19–42.

Borjas, B., 1987. Self-selection and the earnings of immigrants. American Economic Review 77 (4), 531–553.

Chiquiar, D., Hanson, G., 2005. International migration, self-selection, and the distribution of wages: evidence from Mexico and the United States. Journal of Political Economy 113, 239–281.

Co, K., Gang, I., Yum, M.-S., 2000. Returns to returning. Journal of Population Economics 13, 57–79.

Commander, S., Kangasniemi, M., Winters, L., 2003. The brain drain: curse or boom? IZA discussion paper 809.

ETF, 2008. The contribution of human resources development to migration policy in Ukraine. European Training Foundation country report: Ukraine.

Friedberg, R., 2000. You can't take it with you? Immigrant assimilation and the portability of human capital. Journal of Labour Economics 18 (2), 221–251.

Gibson, J., McKenzie, D., 2012. The economic consequences of brain drain of the best and brightest: microeconomics evidence from five countries. The Economic Journal 122 (560), 339–375.

Greene, W. H., 2012. Econometric analysis, 7th Edition. Prentice hall.

Grogger, J., Hanson, G. H., 2011. Income maximization and the selection and sorting of international migrants. Journal of Development Economics 95 (1), 42–57.

Hanson, G. H., 2006. Illegal migration from Mexico to the United States. Journal of Economic Literature 44 (4), 869–924.

Heckman, J., Rubinstein, Y., 2000. The importance of noncognitive skills: Lessons from the GED testing program. American Economic Review 91 (2), 145–149.

Ioannides, Y. M., Loury, L. D., 2004. Job information networks, neighborhood effects, and inequality. Journal of Economic Literature 42 (4), 1056–1093.

Jackson, M., 2008. Social and Economic Networks. Princeton university Press.

Kostenko, W., Harris, M., Zhao, X., 2012. Occupational transition and country-of-origin effects in the early stage occupational assimilation of immigrants: some evidence from Australia. Applied Economics 44 (31), 4019–4035.

Libanova, E. M., 2009. External labor migration of the Ukrainian population (in Ukrainian). Ukrainian Center for Social Reforms, Ukrainian Statistical Committee, Kiev.

Luca, G. D., 2008. SNP and SML estimation of univariate and bivariate binary-choice models. Stata Journal 8 (2), 190–220.

Luca, G. D., Perotti, V., 2011. Estimation of ordered response models with sample selection. Stata Journal 11 (2), 213–239.

Mahmoud, T. O., Trebesch, C., 2010. The economics of human trafficking and labour migration: Micro-evidence from Eastern Europe. Journal of Comparative Economics 38 (2), 173–188.

Martin, R., Radu, D., 2012. Return migration: The experience of Eastern Europe. International Migration 50 (6), 109–128.

Martins, M. F. O., 2001. Parametric and semiparametric estimation of sample selection models: an empirical application to the female labour force in Portugal. Journal of Applied Econometrics 16 (1), 23–39.

Mattoo, A., Neagu, I. C., Özden, Ç., 2008. Brain waste? Educated immigrants in the US labor market. Journal of Development Economics 87, 255–269.

Mayda, A. M., 2010. International migration: a panel data analysis of the determinants of bilateral flows. Journal of Population Economics 23 (4), 1249–1274.

McKenzie, D., Rapoport, H., 2010. Self-selection patterns in Mexico–U.S. migration: the role of migration networks. Review of Economics and Statistics 92 (4), 811–821.

Moraga, J. F.-H., 2011. New evidence on emigrant selection. Review of Economics and Statistics 93 (1), 72–96.

Munshi, K., 2003. Networks in the modern economy: Mexican migrants in the US labor market. The Quarterly Journal of Economics 118 (2), 549–599.

Pedersen, P., Pytlikova, M., Smith, N., 2008. Selection and network effects – migration flows into OECD countries 1990–2000. European Economic Review 52 (7), 1160–1186.

Rooth, D.-O., Saarela, J., 2007. Selection in migration and return migration: Evidence from micro data. Economic Letters 94, 90–95.

Turner, T., 2010. The jobs immigrants do: issues of displacement and marginalisation in the Irish labour market. Work, Employment and Society 24 (2), 318–336.

Van de Ven, W., Pragg, B. V., 1981. The demand for deductibles in private health insurance: A probit model with sample selection. Journal of Econometrics 17 (2), 229–252.

Appendix

C.1 Tables

Table C.1.1: Emigration rates and chosen destinations by region in Ukraine.

ID	Region	Emigration rate, %	Chosen destination, % Russia	EU27	ROW
1	Zhytomyr	1.8	36.8	63.2	0.0
2	Zaporizhia	3.4	20.1	35.4	44.5
3	Kharkiv	5.2	62.7	26.9	10.4
4	Donetsk	5.7	43.5	52.3	4.2
5	Vinnytsia	5.9	0.0	39.7	60.3
6	Mykolaiv	6.2	25.7	65.0	9.4
7	Kirovograd	6.2	50.0	75.0	25.0
8	Krym	6.7	47.4	25.9	26.7
9	Cherkasy	6.9	40.2	48.9	10.9
10	Poltava	7.1	51.7	48.3	0.0
11	Kyiv	7.2	21.0	69.7	9.2
12	Khmelnytskyi	8.2	29.0	71.0	0.0
13	Chernigiv	9.5	27.3	43.6	29.1
14	Chernivtsi	10.5	0.0	84.4	15.6
15	Lviv	11.6	23.3	75.2	1.5
16	Zakarpattia	11.8	38.2	55.7	6.1
17	Dnipropetrovsk	11.8	68.4	20.4	11.2
18	Sumy	14.3	75.1	13.0	11.9
19	Kherson	14.5	35.3	56.0	8.7
20	Lugansk	15.1	89.7	7.6	2.7
21	Ivano-Frankivsk	17.1	19.7	65.6	14.8
22	Volyn	17.9	49.7	50.3	0.0
23	Odesa	22.3	25.6	29.4	45.0
24	Rivne	37.6	44.4	52.2	3.4
25	Ternopil	41.9	9.4	70.7	19.9

Table C.1.2: t-test for the differences of means across migration categories.

Variable	CA vs RM			CA vs PM			CA vs NM			RM vs PM			RM vs NM			PM vs NM		
Male	-0.07	**	(0.04)	0.04		(0.04)	0.22	***	(0.03)	0.11	***	(0.04)	0.29	***	(0.03)	0.18	***	(0.02)
Age:	-0.48		(0.89)	5.83	***	(0.88)	-0.71		(0.68)	6.30	***	(0.92)	-0.23		(0.8)	-6.53	***	(0.73)
15-19	base category																	
20-29	-0.09		(0.14)	0.37	***	(0.07)	0.07	***	(0.01)	0.33	***	(0.07)	0.05	***	(0.01)	-0.01		(0.01)
30-39	-0.14		(0.15)	0.46	***	(0.07)	0.05	***	(0.01)	0.48	***	(0.08)	0.04	***	(0.01)	-0.05	***	(0.01)
40-49	-0.18		(0.15)	0.47	***	(0.07)	0.05	***	(0.01)	0.53	***	(0.07)	0.06	***	(0.01)	-0.05	***	(0.01)
50-59	-0.04		(0.15)	0.57	***	(0.08)	0.03	***	(0.01)	0.49	***	(0.09)	0.02	*	(0.01)	-0.07	***	(0.01)
Marital status:																		
Single	base category																	
Married	-0.10	**	(0.05)	0.26	***	(0.04)	0.00		(0.01)	0.35	***	(0.04)	0.02	***	(0.00)	-0.06	***	(0.00)
Cohabitation	-0.18	*	(0.1	0.12		(0.1)	-0.01		(0.02)	0.28	***	(0.09)	0.02		(0.01)	-0.04	***	(0.01)
Divorced	-0.26	***	(0.08)	0.04		(0.08)	-0.03	**	(0.01)	0.28	***	(0.07)	0.02	*	(0.01)	-0.05	***	(0.01)
Widowed	-0.26		(0.19)	0.02		(0.18)	-0.05	***	(0.02)	0.26		(0.17)	-0.01		(0.01)	-0.08	***	(0.02)
Education:																		
Secondary and below	base category																	
Vocational	-0.27	***	(0.07)	0.19	***	(0.06)	0.02	**	(0.01)	0.43	***	(0.06)	0.06	***	(0.00)	-0.01	*	(0.00)
Higher	-0.04		(0.07)	0.17	***	(0.06)	0.04	***	(0.01)	0.18	***	(0.06)	0.02	***	(0.00)	0.00		(0.00)
Obs.	635			688			6090			585			5987			6040		

Notes: Each estimate is a marginal effect from a probit regression on a respective group of dummy variables. "CA" stands for currently abroad, "RM" denotes Return migrants, "PM" means prospective migrants and "NM" is for non-migrants. Robust standard errors are in parentheses. *** - 1%, ** - 5%, * - 10% significance levels.

Table C.1.3: t-test for the differences of means between males and females.

Variable	Currently broad			Returnees			Prospective			Non-migrants		
Age:	0.19		(1.26)	-1.16		(1.44)	-0.59		(1.35)	-1.42	***	(0.34)
15-19	base category											
20-29	0.04		(0.17)	0.08		(0.24)	-0.02		(0.08)	-0.03		(0.02)
30-39	0.06		(0.17)	0.10		(0.24)	0.02		(0.09)	-0.05	*	(0.03)
40-49	0.16		(0.17)	0.09		(0.23)	0.03		(0.09)	-0.07	***	(0.03)
50-59	-0.05		(0.17)	-0.03		(0.24)	-0.06		(0.11)	-0.08	***	(0.03)
Marital status:												
Single	base category											
Married	0.09		(0.06)	-0.02		(0.07)	0.09		(0.06)	-0.10	***	(0.02)
Cohabitation	-0.04		(0.13)	-0.05		(0.13)	-0.19		(0.14)	-0.10	***	(0.03)
Divorced	-0.18		(0.11)	-0.44	***	(0.1)	-0.12		(0.1)	-0.21	***	(0.02)
Widowed	-0.62	**	(0.28)	-0.21		(0.22)	-0.61	***	(0.24)	-0.36	***	(0.04)
Education:												
Secondary	base category											
Vocational	-0.01		(0.09)	0.08		(0.11)	0.11		(0.09)	0.05	***	(0.02)
Higher	-0.16	**	(0.08)	-0.08		(0.11)	-0.15	*	(0.07)	-0.08	***	(0.02)
Obs.	369			266			319			5721		

Notes: Each estimate is a marginal effect from a probit regression (1 for males) on a respective group of dummy variables. Robust standard errors are in parentheses. *** - 1%, ** - 5%, * - 10% significance levels.

Table C.1.4: Estimated average marginal effects from the log-likelihood function (3.7).

	Selection equation						Outcome equation					
	Baseline			Alternative			Baseline			Alternative		
	dy/dx	SE		dy/dx	SE		dy/dx	SE		dy/dx	SE	
reg_north	base category			base category			base category			base category		
reg_west	0.068	0.019	***	0.068	0.019	***	0.011	0.023		0.010	0.011	
reg_east	0.027	0.019		0.027	0.019		0.010	0.023		0.007	0.010	
reg_south	0.079	0.018	***	0.077	0.018	***	-0.004	0.022		0.015	0.010	
town_50	base category			base category			base category			base category		
town_100	-0.056	0.021	***	-0.057	0.021	***	-0.011	0.026		-0.014	0.011	
town_200	0.013	0.016		0.013	0.016		-0.009	0.020		-0.006	0.008	
town_500	0.007	0.025		0.009	0.025		0.049	0.029	*	0.013	0.011	
town_1000	-0.031	0.016	*	-0.031	0.016	*	-0.066	0.024	***	-0.025	0.009	***
male	0.110	0.019	***	0.110	0.019	***	0.039	0.018	**	0.020	0.009	**
age	0.000	0.001		0.000	0.001		0.000	0.001		0.001	0.000	*
family_below	base category			base category			base category			base category		
family_above	0.127	0.026	***	0.129	0.026	***	0.027	0.031		0.001	0.015	
family_avg	0.050	0.016	***	0.051	0.016	***	0.032	0.018	*	0.003	0.008	
single	-0.073	0.020	***	-0.074	0.020	***	-0.039	0.018	**	-0.022	0.009	***
hh_other	base category			base category			base category			base category		
hh_head	-0.087	0.023	***	-0.086	0.023	***	-0.024	0.020		-0.006	0.009	
hh_spouse	-0.119	0.028	***	-0.118	0.029	***	-0.022	0.025		-0.019	0.011	
hh_small	base category			base category			base category			base category		
hh_med	-0.002	0.016		-0.002	0.016		-0.012	0.016		-0.006	0.007	
hh_large	0.014	0.023		0.011	0.023		0.014	0.026		-0.007	0.010	
educ1	base category			base category			base category			base category		
educ2	0.113	0.026	***	0.113	0.026	***	0.032	0.027		0.039	0.017	**
educ3	0.109	0.023	***	0.109	0.024	***	0.141	0.037	***	0.076	0.019	***

continued on the next page

Notes: Standard errors are clustered by $family\_id$. *** - 1%, ** - 5%, * - 10% significance levels.

continuation of Table C.1.4

	Selection equation						Outcome equation					
	Baseline			Alternative			Baseline			Alternative		
	dy/dx	SE		dy/dx	SE		dy/dx	SE		dy/dx	SE	
shift_ua							0.140	0.041	***	-0.025	0.007	***
times_traveled							0.005	0.004		0.001	0.002	
ROW							base category			base category		
FSU							0.010	0.024		0.034	0.014	**
EU							0.013	0.023		0.031	0.013	**
USA							0.035	0.038		0.029	0.017	*
stay_duration							0.000	0.000		0.000	0.000	*
status_other							base category			base category		
status_work							-0.016	0.020		-0.007	0.008	
status_res							-0.053	0.035		-0.004	0.012	
reason_pay							0.044	0.021	**	0.008	0.008	
language							-0.061	0.031	*	-0.026	0.009	***
migraid_rec							0.010	0.016		0.006	0.007	
sponsor_nat_other							base category			base category		
sponsor_nat_ua							-0.020	0.032		-0.036	0.016	**
sponsor_nat_abr							-0.008	0.018		0.004	0.008	
dest_fam							-0.004	0.028		-0.002	0.012	
shift_fam							0.607	0.234	***	0.230	0.048	***
employed							base category			base category		
unempl_y_ua	0.036	0.024		0.038	0.025							
unempl_o_ua	0.115	0.022	***	0.125	0.023	***						
migr_fam	0.114	0.026	***	0.114	0.026	***						
athrho	1.541	0.423	***	1.510	0.283	***						
ln(L)	-2035.67			-1986.28								
Censored obs.	5721											
Uncensored obs.	614											

Table C.1.5: Estimated average marginal effects of the semi-nonparametric estimator of Luca (2008) and Luca and Perotti (2011).

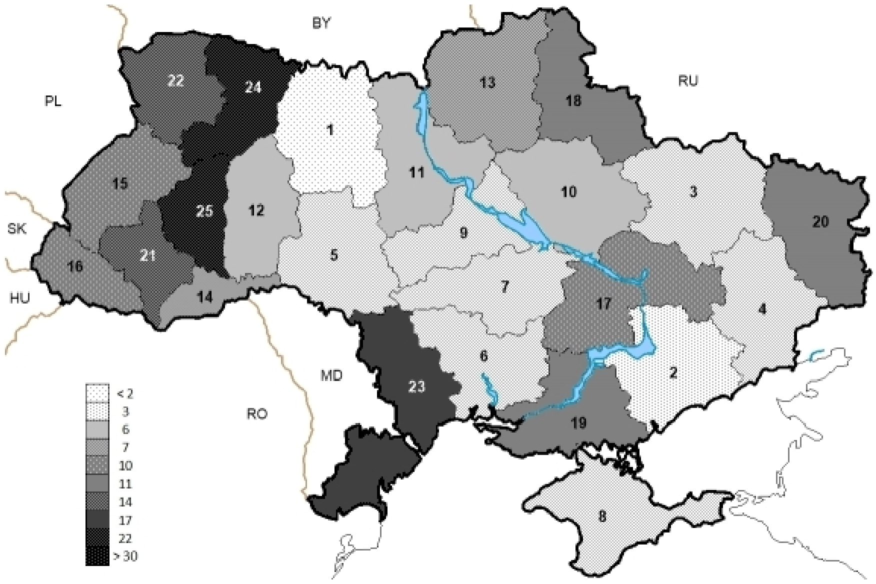
	Selection equation						Outcome equation					
	Baseline			Alternative			Baseline			Alternative		
	dy/dx	SE		dy/dx	SE		dy/dx	SE		dy/dx	SE	
reg_west	0.073	0.074		0.072	0.024	***	-0.001	0.066		0.004	0.061	
reg_east	0.029	0.098		0.035	0.024		0.024	0.218		0.041	0.055	
reg_south	0.078	0.057		0.084	0.023	***	-0.051	0.036		0.020	0.068	
town_100	-0.076	0.069		-0.069	0.028	**	0.016	0.347		-0.081	0.073	
town_200	0.002	0.035		0.010	0.019		-0.040	0.132		-0.099	0.052	*
town_500	-0.003	0.025		0.003	0.027		0.157	0.312		-0.019	0.052	
town_1000	-0.039	0.043		-0.033	0.020	**	-0.115	0.218		-0.174	0.054	***
male	0.110	0.089		0.137	0.032	***	-0.044	0.108		-0.013	0.037	
age	-0.001	0.001		0.000	0.001		0.001	0.004		0.005	0.002	**
family_above	0.120	0.099		0.133	0.029	***	-0.052	0.217		-0.081	0.062	
family_avg	0.038	0.072		0.054	0.020	***	0.026	0.243		-0.007	0.033	
single	-0.114	0.078		-0.101	0.030	***	-0.039	0.062		-0.128	0.064	**
hh_head	-0.125	0.056	**	-0.111	0.033	***	0.011	0.035		0.007	0.065	
hh_spouse	-0.172	0.101	*	-0.133	0.033	***	0.030	0.063		0.007	0.059	
hh_med	-0.021	0.019		-0.008	0.018		0.011	0.026		-0.043	0.038	
hh_large	-0.006	0.034		0.008	0.026		0.052	0.125		-0.083	0.064	
educ2	0.098	0.096		0.121	0.029	***	-0.007	0.372		0.118	0.091	
educ3	0.084	0.026	***	0.109	0.026	***	0.264	0.279		0.360	0.125	***
shift_ua							0.384	0.099	***	-0.178	0.060	***
times_traveled							0.017	0.013		0.009	0.006	
FSU							0.050	0.090		0.264	0.098	***
EU							0.037	0.061		0.233	0.090	***
USA							0.109	0.103		0.203	0.102	**
stay_duration							0.000	0.001		-0.001	0.000	**
status_work							-0.071	0.083		-0.088	0.047	*
status_residence							-0.174	0.178		-0.057	0.063	
reason_pay							0.085	0.119		0.037	0.037	
language							-0.185	0.320		-0.162	0.057	***
migraid_rec							0.043	0.170		0.063	0.030	**
sponsor_nat_ua							-0.002	0.233		-0.212	0.087	**
sponsor_nat_abr							0.033	0.095		-0.008	0.036	
dest_fam							-0.108	0.044	**	-0.122	0.066	*
shift_fam							1.194	0.652	*	1.576	0.427	***
unempl_y_ua	-0.003	0.027		0.025	0.031							
unempl_o_ua	0.111	0.041	***	0.122	0.027	***						
migr_fam	0.117	0.022		0.125	0.021	***						
rho	0.045			-0.003								
ln(L)	-2031.16			-1975.79								
Censored obs.	5721											
Uncensored obs.	614											

Notes: Standard errors are clustered by $family\_id$. *** - 1%, ** - 5%, * - 10% significance levels.

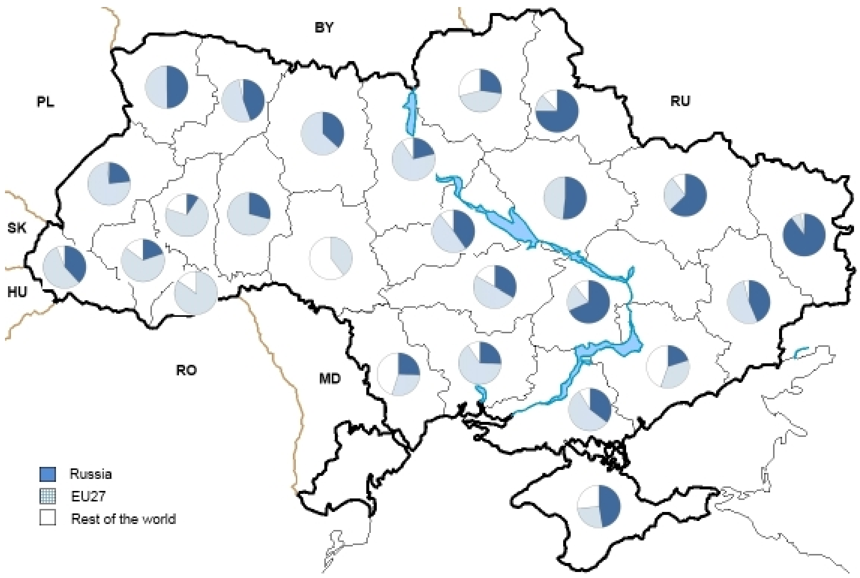
Table C.1.6: Definitions of covariates.

Variable	Definition
shift_ua	$= 1$ if the respondent downshifted in Ukraine, and 0 otherwise.
reg_west reg_east reg_south	Set of dummy variables, $= 1$ if the respondent comes from the West, East or South of Ukraine respectively, and 0 otherwise. The base region is the North.
town_100 town_200 town_500 town_1000	Set of dummy variables, $= 1$ if the respondent lives in a settlement with population $[100k - 200k)$, $[200k - 500k)$, $[500k - 1000k)$ and $\geq 1000k$ people respectively. The base category is $[50k - 100k)$.
male	$= 1$ if the respondent is male, and 0 otherwise.
age	respondent's age in years.
family_below family_avg family_above	Set of dummy variables, $= 1$ if the respondent's self-reported income is below average, average or above average respectively.
single	$= 1$ if the respondent is single, divorced or widowed, and 0 otherwise.
hh_head hh_spouse	Set of dummy variables, $= 1$ if the respondent is the head of the household or the spouse of the head respectively, and 0 otherwise. The base category is all others (son, daughter *etc*).
hh_med hh_large	$= 1$ if the respondent comes from a medium-sized (3 or 4 members) or large (5 and above) household, and 0 otherwise. The base group is small households with at most two members.
times_traveled	number of times the respondent traveled to the same country for the same purpose within the last three years (excluding occasional returns to Ukraine).
FSU EU USA	Destination dummy variables, $= 1$ if the respondent went to a country of the former Soviet Union (excluding the Baltic countries), EU27 or North America respectively, and 0 otherwise. The rest of the world is the base.
stay_duration	Duration of stay (in years) in the destination country.
status_work status_residence	$= 1$ if the respondent has a work permit or permanent residency respectively, and 0 otherwise. The base is all other categories.
reason_pay	$= 1$ if the respondent's primary reason for migration was higher wage or better employment opportunities, and 0 otherwise.
language	$= 1$ if the respondent speaks the language of the destination country or English of an intermediate level or above.
migraid_rec	$= 1$ if the respondent received any help to emigrate from friends / relatives or co-workers in Ukraine, and 0 otherwise.
sponsor_nat_ua sponsor_nat_abr	Set of dummy variable, $= 1$ if the nationality of the respondent's employer / sponsor is Ukrainian or that of the destination country respectively, and 0 otherwise. The base is all other nationalities.
educ1, educ2, educ3	Dummy variable for secondary education, vocational training, and higher education (Bachelor's degree and above) respectively.
unempl_y_ua unempl_o_ua	Set of dummy variables, $= 1$ if the respondent is unemployed and aged $[15 - 30]$ or $(30 - 59]$ respectively, and 0 otherwise.
migr_fam	$= 1$ if the family has a migrant (besides the current one), and 0 otherwise.

C.2 Figures

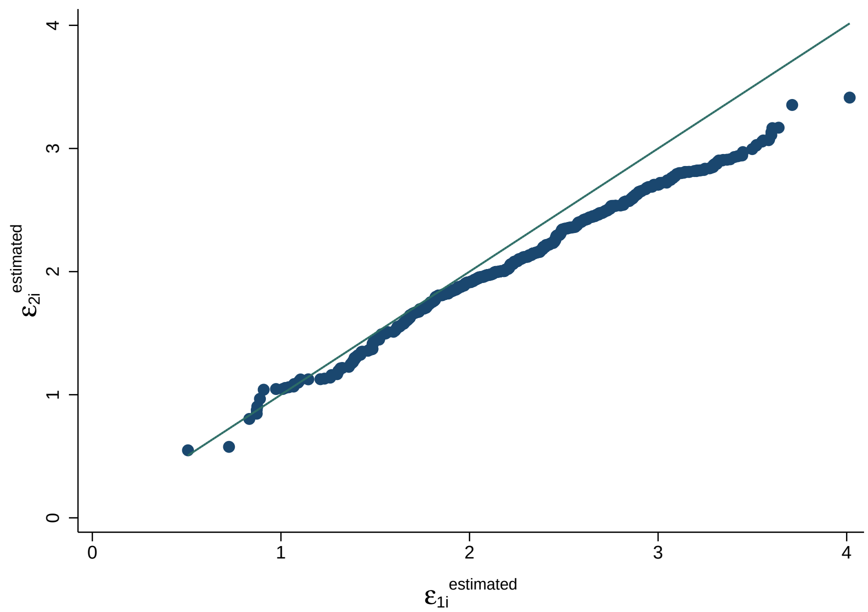


(a) *Emigration rates, %.*

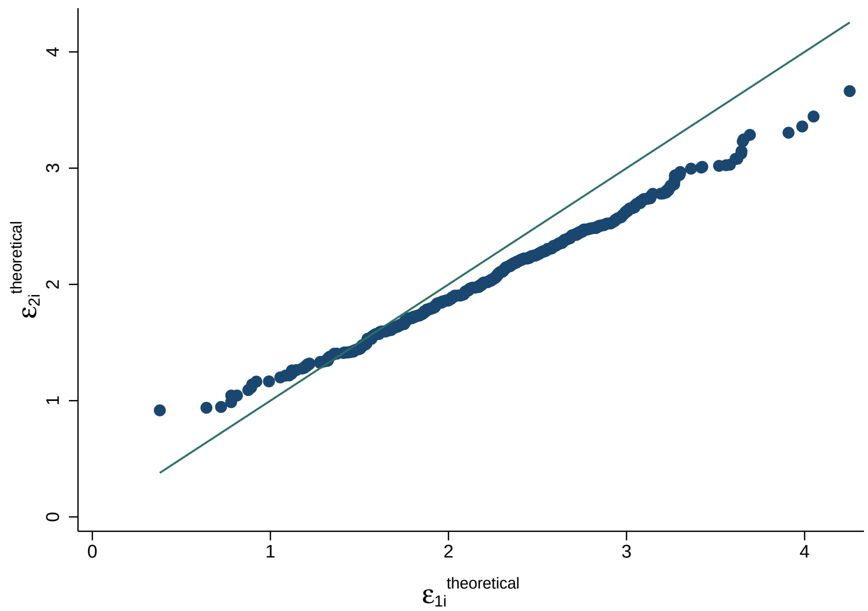


(b) *Chosen destinations.*

Figure C.2.1: Emigration rates and chosen destinations.



(a) *Estimated $\hat{\epsilon}_{1i}$ and $\hat{\epsilon}_{2i}$.*



(b) *Theoretical bivariate normal.*

Figure C.2.2: Quantile-quantile plots.

Chapter 4

Commuting Patterns of Czech Households Exposed to Flood Risk from the River Bečva

Abstract

Using unique data collected in October–December 2012 we estimate the relationship between commuting for work and the level of individual exposure to floods. We find that commuters on average have higher earnings than non-commuters. Individuals affected by one flood commute 11.2% more than unaffected individuals. We conjecture that this increase is linked to intentions to cover flood-related losses, decrease households' vulnerability to flood risk or out-migrate from the risk areas. Individuals affected by at least two floods are 20.2% less likely to commute relative to those unaffected. We explain this nonlinear effect by the fact that many households out-migrate after the first flood. Stayers commute less because they are different from non-stayers in some underlying characteristics related to education, employment and family circumstances, which strongly affect commuting behavior. We further find that in a commuting family an individual is 53.8% more likely to commence commuting relative to a non-commuting family. The choices of commuting destinations are often similar to those of other family members.

Keywords: commuting; income gap; flood risk; selection; adaptation; probit

JEL classification: Q01; Q50; Q56

This paper is a joint work with Robert Stojanov, Barbora Duží, and David Juřička. The earlier version was published in Environmental Hazards, 2014, 13(1), pp. 58–72. We are thankful to Oliver Bakewell, Jeanette Schade, and Thomas Faist for their comments during the conference at Bielefeld University (2012). The financial support of grants CZ.1.05/1.1.00/02.0073, CZ.1.07/2.4.00/31.0056, LD 13032 and RVO 67179843 is gratefully acknowledged.

4.1 Introduction

Commuting for work is an economic decision. The benefits of commuting come from the higher income that a commuter and his family can enjoy. The costs of commuting are related to less leisure and less time spent with the family. It is therefore interesting to know how various individuals weigh these benefits and costs.

One has to separate two distinct but interrelated decisions: a decision to commute (extensive margin) and how far to commute (intensive margin). In this paper, we research these two decisions in a unified framework. We postulate two research questions: "What is the character of the relationship between commuting for work and the level of household exposure to floods?" and "Do individuals from flood affected households commute shorter or longer distances compared to those from unaffected ones?"

In our survey most individuals report a significant wage gap between what they earn in destination cities and their current places of residence. Wages are endogenous in that they are determined by many factors which are often not entirely observed. One cannot expect much economic activity in regions that are frequently exposed to floods. Therefore, exposure to floods is an important determinant of economic activity, which affects the likelihood of commuting. This simple intuition explains why we think of commuting as an activity strongly related to individual flood experience.

Christensen and Christensen (2003) and Kundzewicz et al. (2013) indicate that Central Europe was severely affected by repeated floods within the last two decades. Borga et al. (2011) emphasize the necessity for more empirical research to fill in the data gap on flood evidence from small river basins and their effects on local residents. Our research contributes significantly to filling in this gap – we collect an individual level dataset which allows the investigation of the effects of floods on economic activity of households and their members. To our best knowledge this paper is one of the few attempts to systematically analyze floods from the River Bečva, its smaller water sources and their effects on the economic behavior of affected residents.

This research is linked to the literature on the internal mobility of people caused by environmental factors. Many authors find that floods are devastating for the well-being of affected communities and companies (see Kreibich and Thieken, 2009; Yeo, 2002; Kreibich et al., 2007 and Thieken et al., 2007). When environmental risks are relatively high, the lives of local residents are affected in a way that they have to undergo adaptation to increase their resilience. This includes the acquisition of various equipment (Grothmann and Reusswig, 2006 and Wachinger et al., 2013), purchase of insurance (Botzen et al., 2009), internal mobility or emigration. Barrios et al. (2006) finds that the decreasing amounts of rainfall increased the rates of urbanization in the countries of sub-Saharan Africa. Reuveny (2007) mentions that repeated floods are often the reason behind migration within and from Bangladesh. Dust storms caused the decline of agriculture in the US Great Plains followed by emigration. Warner (2010) provides evidence of environmentally induced migration in Mozambique, Vietnam, and Egypt.

Our study focuses on rural regions in the central part of the Bečva river in the Eastern part of the Czech Republic. Individuals there live in villages or small towns and usually commute for work to nearby larger towns or cities. Our main finding is the existence of a nonlinear effect of floods on the patterns of commuting of affected respondents. Individuals affected by only one flood are 11.2% more likely to commute, whereas individuals affected by at least two floods are 20.2% less likely to do so. We explain this finding by the fact that the affected individuals commute, in part, to accumulate financial resources to cover damages from floods and decrease households' vulnerability to flood risk. Some households manage to out-migrate permanently from risk areas. Those who remain and get exposed to the second flood are those who, for some reasons, were not active commuters after the first flood. These reasons are complex and are related to education, employment details, family circumstances and individual mobility costs. These findings are supported by qualitative information from households that we collected during face-to-face interviews.

We also find support for the "network effect" hypothesis, according to which an average respondent with an active commuter in the family is 53% more likely to engage in commuting. New commuters who already have a long-distance commuter in the family are also likely to commute long distance.

This paper is structured as follows. We first describe the sampling frame and survey instrument. Then, we provide descriptive statistics on respondents in the sample. Further, we formulate an econometric model, estimate it, interpret the results and conclude.

4.2 Survey design

The population of interest consists of households residing in the risk areas of the River Bečva in the Eastern part of the Czech Republic.[1] We stratify the population with respect to administrative regions and the level of past exposure to floods: badly affected areas (occurrence of at least two floods), moderately affected areas (occurrence of one flood) and unaffected areas (no floods occurred and location within 200 meters from the moderately affected area). Data on the distribution of houses across the three risk areas are taken from CHMI (2012). We distribute the total number of interviews proportionally to population in each stratum. Figure D.1.1 illustrates the population of interest and location of houses across risk zones on the example of Choryně and Poličná regions. The distribution of interviews across regions is shown in Table 1.

Table 1: Distribution of observations across administrative regions.

	Households		Individuals	
Admin. region	N	%	N	%
Choryně	30	9.9	84	9.6
Hrachovec	28	9.2	92	10.6
Hustopeče nad Bečvou	12	3.9	32	3.6
Juřinka	14	4.6	33	3.8
Krhová	31	10.2	84	9.6
Lhotka nad Bečvou	18	5.9	52	5.9
Milotice nad Bečvou	10	3.3	30	3.4
Poličná	32	10.5	91	10.4
Střítež nad Bečvou	29	9.6	85	9.7
Ústí	31	10.2	96	11.0
Zašová	31	10.2	76	8.7
Zubří	38	12.5	120	13.7
Total	304		875	

The survey instrument consists of two parts: household and individual level questions. The household level questions are aimed at learning about past flood experience, responses during recovery phase, insurance and preparedness for potential floods. The individual level questions intend to learn the demographic characteristics, details of economic activity and migration / mobility intentions of each adult member of a households.

[1] In a related study, Brázdil et al. (2011) research the River Morava, the main stem for the River Bečva, which remained rather untouched in their analysis.

The questionnaire consists of many open-ended questions, in which each respondent can evaluate their household's vulnerability to flood risk and express their opinions on the effectiveness of government anti-flood measures. These questions help us better understand circumstances of the surveyed households.

4.3 Descriptive statistics

The collected sample contains data on 304 households and 875 individuals over five flood occurrences: 1997, 2002, 2006, 2009 and 2010. In line with official data our research finds (see Table 2) that the most severe flood took place in 1997 and affected 184 households and 568 individuals in the collected sample. All subsequent floods were less severe. Slightly more than one third of all households had experience with only one flood, 28.3% experienced two floods and 8.2% of the surveyed households experienced at least three floods.

Table 2: Flood occurrences and cumulative flood experience.

	Households		Individuals	
Year	N	%	N	%
1997	184	60.5	568	64.9
2002	37	12.2	123	14.1
2006	23	7.6	66	7.5
2009	57	18.8	160	18.3
2010	66	21.7	193	22.1
Cumulative flood experience:				
One flood	108	35.5	303	34.6
Two floods	86	28.3	262	29.9
At least three floods	25	8.2	79	9

Table 3 provides data on self-reported losses from floods. Throughout all five floods most households suffered up to CZK 50k (EUR 2k) in losses, which suggests the persistent but not devastating nature of the floods. We have two reasons to believe that the reported losses might be slightly mismeasured. Firstly, in a few cases respondents had difficulty quantifying losses because their damaged houses were never fixed after the flood(s). Secondly, insurance companies often participated in fixing affected houses or replacing damaged equipment. In such cases respondents could not give reliable estimates of the value of goods delivered to them by insurance companies.

Table 3: Reported financial losses.

in *CZK* − >	0–50k	50k–100k	100k–200k	200k–500k	500k–1 mln
in *EUR* − >	0–2k	2k–4k	4k–8k	8k–20k	20k–40k
1997	121	26	13	13	5
2002	29	3	.	.	1
2006	5	4	1	3	1
2009	37	6	6	.	.
2010	55	7	4	.	.

Note: A dot means zero value.

Prior to each flood three fourths of households had insurance contracts. The remaining one fourth did not have insurance because either it was too expensive to purchase or no insurance company would agree to insure their houses located in high risk areas. Table 4 shows the distribution of insurance settlements across affected households. The settlements are shown as the shares of reported losses covered by insurance companies.

Table 4: Number of households that had a given share of the losses covered by insurance.

Year	10%	20%	30%	40%	50%	60%	70%	80%	90%
1997	89	82.4	75.8	69.2	20.9	19.8	16.5	9.9	8.8
2002	88.9	77.8	77.8	77.8	33.3	22.2	22.2	22.2	22.2
2006	100	83.3	50	50	33.3	33.3	16.7	16.7	16.7
2009	94.1	88.2	70.6	58.8	17.6	17.6	5.9	5.9	5.9
2010	100	90	90	86.7	50	50	46.7	33.3	30

Interestingly, after the flood in 1997 insurance covered at least 40% of losses to 69.2% of affected households. However, at least 50% were covered to only 20.9% of households. Data show that for all five floods insurance companies were unwilling or unable to cover more than 50% of losses for the vast majority of households. We have two explanations for that. The first is related to how insurance companies operate: after the flood many customers claimed losses, so had insurance companies been generous in payments, many of them would have gone bankrupt. Indeed, several local insurance companies stopped operating after the flood in 1997. The second explanation concerns the nature of insurance contracts. Many respondents were under-insured, in that their contracts covered fewer assets than they thought. After the floods they claimed losses for assets that had not been insured.[2]

Basic demographic characteristics are provided in Table 5. We have almost equal shares of males and females, most of whom (62%) are married, 23.2% are single, 9.5% are widowed and 4.2% are divorced. 40% of respondents have completed secondary education, slightly less, 34.6%, have incomplete secondary education and only 9,6% have a Master's degree or above.

Table 5: Basic demographic characteristics.

	N	%		N	%
Male	439	50.2	Occupation type:		
Marital status:			Low-skilled	136	15.8
Single	203	23.5	Medium-skilled	159	18.5
Married	542	62.7	High-skilled	64	7.4
Divorced	37	4.3	Entrepreneur	45	5.2
Widowed	83	9.5	Retired	333	38.7
Education:			Student	57	6.6
Primary	101	11.6	Maternity leave	25	2.9
Incomplete secondary	302	34.7	Unemployed	42	4.9
Complete secondary	357	41.0			
Professional	12	1.4	Commuters:	267	65.1
Bachelor's degree	15	1.7			
Master's degree and above	84	9.6			

In the questionnaire, we developed a scale to rank the skill intensity of employment occupations. The distribution of respondents across low-, medium- and high-skilled occupations is 15.8%, 18.5% and 7.4% respectively. There are 333 retirees, 57 students, 42 unemployed and 25 women on maternity leave in the sample. The share of commuters (out of the pool of working age subsample excluding unemployed, students and women on the maternity leave) is 65.1%, or 267 individuals.

[2] A typical insurance contract has separate provisions for insuring a house (walls, doors, cellar, *etc.*) and assets in the house (boiler, furniture, electronics, *etc.*).

4.4 Estimation

4.4.1 Mincerian wage regression

To research the determinants of earnings we have to estimate the Mincerian wage regression on the subsample of working age individuals who reported their income. We exclude pensioners, women on maternity leave, students and unemployed from the estimation. We do not observe income for a significant share of working individuals due to non-reporting (response rate to the income question is 46%). It is therefore not convincing to rely on the estimates of income gap that come from a truncated distribution.

There are reasons to suspect that the non-reporting of income happens on a systematic (non-random) basis and is related to the true level of income. To account for that we use the Heckman selection procedure (Heckman, 1979) and write a selection equation for whether a respondent reports the income of each family member.[3] As an instrument for reporting income we use the number of working adults in the family. The intuition is that for a respondent from a large family it takes more time to report employment details about each member of the family. Thus, such respondents are more likely to say that they do not know or opt out of an interview completely.[4]

The wage equation can be written as follows:

$$E[ln(wage_i)] = X_i^{'}\delta_1 + Z_i^{'}\delta_2 + D_i^{'}\delta_3 + E(\epsilon_{1i}|\epsilon_{0i} > -X_{i0}^{'}\delta_0), \tag{4.1}$$

where $wage_i$ is the reported wage of individual i, X_i is a column vector of individual characteristics that include gender, age, family status and the number of children. Z_i is a column vector that contains education, experience, a dummy variable for whether person i commutes and occupation type dummy variables. $D_i^{'}$ are region fixed effects to account for regional heterogeneity in average incomes. The last term accounts for the fact that the income for some individuals is not reported and the observed income distribution is truncated. Vector X_{i0} contains the same covariates as X_i plus a variable for the number of working adults in a family. This variable serves as an instrument to predict respondent's

[3] An alternative procedure would be to construct a likelihood function in the spirit of the Tobit model.

[4] This instrument indeed has a significant predictive power. If family size increases by one individual, the probability of reporting income drops by 5.2%. This estimate is significant at 5%.

decision to report income.

Equation (4.1) is estimable with OLS, because the term $E(\epsilon_{1i}|\epsilon_{0i} > -X'_{i0}\delta_0)$ can be expressed in a closed form assuming that the error terms have bivariate normal distribution. The exact definitions of the covariates are given in Table D.2.1. The estimates of equation (4.1) are shown in Table 6.

Table 6: OLS estimates of the Mincerian wage regression (4.1).

	Estimate		SE
commute	0.196	***	(0.05)
male	0.248	***	(0.04)
age	0.045	**	(0.02)
age2	-0.001	**	(0.00)
educ2	0.155	***	(0.05)
educ3	0.197	**	(0.09)
exper	0.002		(0.00)
married	0.091		(0.06)
kids1	0.231	*	(0.12)
kids2	0.087		(0.12)
kids3	-0.086		(0.09)
occ_type2	0.085		(0.06)
occ_type3	0.225	***	(0.08)
occ_type4	0.327	***	(0.08)
λ_{report}	0.285		(0.18)
cons	8.140		(0.44)

Notes: The regression includes region (obec) dummy variables. The number of observations is 215 and adj. $R^2 = 0.45$. Standard errors are clustered by $family\_id$. *** - 1%, ** - 5%, * - 10% significance levels.

The signs of the estimates are in line with the predictions of economic theory. Age has a concave shape: earnings increase with age at a declining pace. Respondents with more experience, higher education and those in more skill demanding occupations earn more. Males earn more than females – an established fact of gender wage gap. The estimates suggest that respondents who commute for work to nearby larger cities are paid more than those who work locally. In particular, an average commuter earns 19.6% more than a non-commuter.

4.4.2 Determinants of commuting

Given the fact that commuters are higher earners than non-commuters, we wish to investigate whether selection into commuting is somehow linked to the level of exposure of that household to floods. In attempts to cover financial losses from floods, individuals might wish to look for better paying jobs and thus commence commuting or out-migrate permanently from risk areas. Since out-migration is costly, individuals are more likely to decide to commute because the marginal costs of doing so are lower.

To answer this question it is necessary to properly define the dependent variable. We wish to learn if flood affected individuals commute differently relative to unaffected ones. We identified five large and medium-size floods that occurred as depicted in Figure 1. For a respondent who started commuting at some point between 1997 and 2002 it is important to know if he was exposed to the flood in 1997. In the same fashion, for a respondent who started commuting between 2002 and 2006 it is crucial to know if that respondent was affected by floods that occurred in 2002 and 1997. It is of little informative value to know whether that respondent was affected by floods after he had started commuting. Finally, for somebody who started commuting after 2010 we wish to know if that respondent was affected by any of the five researched floods.

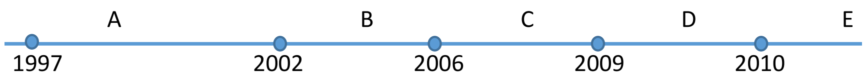


Figure 1: Occurrence of floods.

Based on the described intuition we create three key variables - *commute*, $first\_flood$ and *second_flood*. Variable *commute* equals 1 if a respondent started commuting in any of the five areas - A, B, C, D or E; and 0 otherwise. Dummy variables $first\_flood$ and *second_flood* capture the first and second flood occurrences prior to the start of commuting. If a respondent was affected by all five floods and started commuting between 1997 and 2002, then $commute = 1$, $first\_flood = 1$ and $second\_flood = 0$. If the same individual started commuting between 2002 and 2006, then $commute = 1$, $first\_flood = 0$ and $second\_flood = 1$. It does not help us to know if that respondent was affected by floods in 2006, 2009 and 2010 after he started commuting because this fact does not entail causality. Only floods that occurred prior to the start of commuting could be a contributing factor to the decision to commute. Table 7 depicts the relevance of the created variable. Out of 267 individuals who commuted on the survey date only 146, or 55%, are those for whom the preceding flood occurrence could have been a contributing factor. The remaining 121 individuals commenced commuting prior to the flood date and are not classified as commuters.

Table 7: Discrepancies between commuting on the survey date and the defined commute variable.

		Commute on survey date		
		No	Yes	Total
commute var.	No	126	121	247
	Yes	0	146	146
	Total	126	267	393

To learn the determinants of commuting we estimate the following equation:

$$\begin{aligned} commute_i &= \beta_0 + \beta_1 first\_flood_i + \beta_2 second\_flood_i + \\ &+ \beta_3 loss\_big\_ff_i + \beta_4 loss\_big\_sf_i + \beta_5 cov\_more\_ff_i + \beta_6 cov\_more\_sf_i + \\ &+ \beta_7 educ2_i + \beta_8 educ3_i + \beta_9 gender_i + \beta_{10} age30_i + \beta_{11} age40_i + \\ &+ \beta_{12} age50_i + \beta_{13} married_i + \beta_{14} kids1_i + \beta_{15} kids2_i + \beta_{16} kids3_i + \\ &+ \beta_{17} fam\_com_i + D_i^{'}\theta + \nu_i. \end{aligned} \tag{4.2}$$

Variables $commute$, $first\_flood$ and $second\_flood$ are defined as described above. The second line of equation (4.2) contains dummy variables that describe the level of reported losses and insurance settlements after each of the two floods. The third and fourth lines contain variables that describe individual demographic characteristics. Variable fam_com describes whether the respondent's family already contained a commuter before the start of commuting. With this variable we test the "network effect" hypothesis, which conjectures that it is easier for an individual to start commuting once there is already somebody in the family doing so. D_i is a column vector of region dummies to account for heterogeneity in unobserved region characteristics. The exact definitions of covariates are given in Table D.2.1.

Under the assumption $\nu_i \sim N(0, \sigma^2)$ regression (4.2) is a standard probit model. The estimation results and marginal effects of regression (4.2) are shown in panels one and two in Table 8. Clustering by $family\_id$ accounts for the possibility of correlation of individual error terms within the same household.

The estimation results suggest that the exposure to floods has a sizeable nonlinear effect on the individual probability of commuting. Exposure to the first flood increases the probability of commuting by 11.2% as compared to unaffected individuals (panel two). When the second flood occurs the probability of commuting decreases by 20.2% as compared to unaffected individuals. This nonlinear effect is depicted in Figure 2.

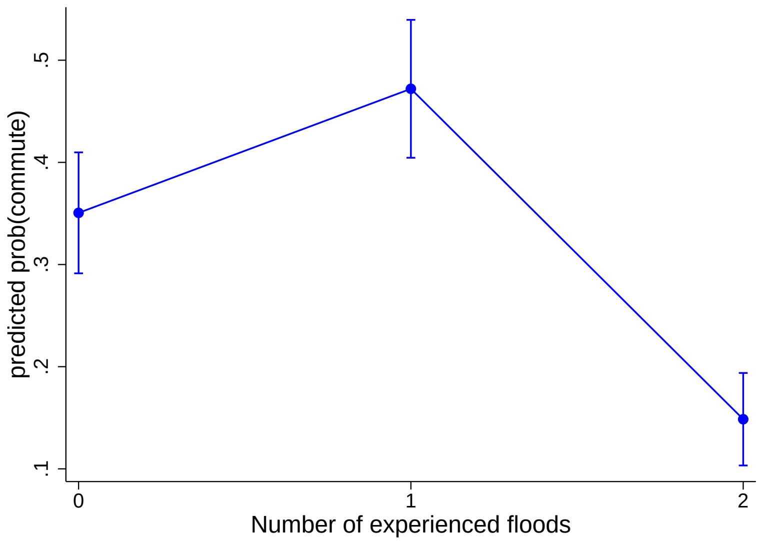


Figure 2: Probability of commuting. Point estimates with 95% confidence intervals. Illustration for the mean values of the variables.

Further, we find that individuals with losses above EUR 2k from any flood are by a slight margin more likely to commute. Individuals for whom insurance settlements exceed 50% of reported losses are roughly 20% less likely to commute.

These estimates point to the fact that commuting for work is indeed related to the intensity of household exposure to floods. Having high losses (above EUR 2k) is a negative shock to households and pushes individuals to look for better employment in large cities to increase the sustainability of their households. Individuals who were sufficiently insured against flood losses (in that the settlements covered at least 50% of reported losses) recovered from the shock more easily than under-insured individuals. The sustainability

Table 8: Probit estimates of equation (4.2).

	Panel one Probit, first flood						Panel two Probit, both floods						Panel three IV, both floods		
	Estimate		SE	dy/dx		SE	Estimate		SE	dy/dx		SE	Estimate		SE
first_flood	1.058	***	(0.22)	0.189	***	(0.04)	0.674	***	(0.24)	0.112	***	(0.04)	0.538		(1.52)
second_flood							-1.220	***	(0.66)	-0.202	***	(0.06)	-1.626	*	(0.95)
loss_big_ff	0.128		(0.36)	0.023		(0.06)	0.083		(0.37)	0.014		(0.06)	0.084		(3.33)
loss_big_sf							0.523		(0.48)	0.087		(0.08)	4.054		(3.95)
cov_more_ff	-1.089	**	(0.44)	-0.195	**	(0.08)	-1.035	**	(0.46)	-0.172	**	(0.07)	-0.995		(1.18)
cov_more_sf							-1.326	**	(0.68)	-0.220	**	(0.11)	-1.281		(1.17)
married	0.111		(0.23)	0.020		(0.04)	0.221		(0.25)	0.037		(0.04)	0.190		(0.38)
male	0.515	***	(0.20)	0.092	***	(0.03)	0.541	**	(0.21)	0.090	***	(0.03)	0.497	*	(0.27)
age30	1.521	***	(0.37)	0.272	***	(0.06)	1.787	***	(0.42)	0.296	***	(0.06)	1.827	***	(0.46)
age40	1.058	***	(0.30)	0.189	***	(0.05)	1.347	***	(0.33)	0.223	***	(0.05)	1.341	***	(0.41)
age50	0.761	***	(0.29)	0.136	***	(0.05)	0.919	***	(0.29)	0.152	***	(0.05)	1.018	***	(0.37)
educ2	-0.276		(0.21	-0.049		(0.04)	-0.381	*	(0.22)	-0.063	*	(0.04)	-0.456		(0.41)
educ3	-0.072		(0.27)	-0.013		(0.05)	0.016		(0.29)	0.003		(0.05)	-0.301		(0.74)
kids1	-0.291		(0.29)	-0.052		(0.05)	-0.613	*	(0.34)	-0.102	*	(0.05)	-0.660	*	(0.34)
kids2	-0.098		(0.28)	-0.017		(0.05)	-0.370		(0.30)	-0.061		(0.05)	-0.287		(0.49)
kids3	-0.536		(0.41)	-0.096		(0.07)	-0.831	*	(0.44)	-0.138	*	(0.07)	-0.675		(0.75)
fam_com	2.686	***	(0.25)	0.481	***	(0.03)	3.246	***	(0.32)	0.538	***	(0.04)	3.346	***	(0.50)
cons	-7.119	***	(0.49)				-7.427	***	(0.62)				-7.622		(464.8)

Notes: Each regression includes region (obec) dummy variables. The number of observations is 378 in each equation. $\ln(L) = -123.2$ and $\ln(L) = -113.6$ in panels one and two, respectively. The Wald chi-squared test in panel three does not reject the null hypothesis of exogeneity (Prob $> \chi^2 = 0.74$). Standard errors are clustered by $family\_id$. *** - 1%, ** - 5%, * - 10% significance levels.

of these households was not compromised by floods.

The findings are in line with qualitative data from respondents. Many of them are unhappy to live in areas of high flood risk and would be glad to out-migrate permanently if only circumstances (in a wide sense) allowed for that. We saw several abandoned houses and learnt from neighbors that their owners had moved out. The houses could not be sold, because they had trivial value on the market. Unfortunately, we did not manage to learn any reliable details about the emigrated households.

Many respondents expressed concern about their insurance contracts, insurance settlements, and rising insurance premiums. Few respondents could not get an insurance contract because an insurance company would not insure a house located in high risk area. For such a household any flood event is a negative shock with which it is left to cope on its own. In most cases insurance companies were parsimonious in settlements. For an under-insured household with low income it means the inability to completely recover from flood losses. Indeed, we saw individuals living in houses still unrepaired several years after the flood(s). To all surveyed households the rising insurance premiums is worrisome. They indicated that the price of an insurance contract doubled over the past decade.

Males, married people, young people, and respondents with at least undergraduate degrees are more likely to commute. Having children is negatively associated with commuting, because individuals substitute their time at work for time with the family. We do not reject the "network effect" hypothesis: an average individual who has somebody already commuting in the family is 53.8% more likely to start commuting.

The variables $first\_flood$, $second\_flood$, $loss\_big\_ff$ and $loss\_big\_sf$ are exogenous to the commute decision, therefore the estimates are unbiased. However, the occurrence of floods and levels of losses are endogenous with respect to location. Houses located on flat slopes closer to the river are more likely to be affected by rising water and have higher losses than houses located on steep slopes. Thus if we find an instrument that predicts house location and does not affect *commute* variable directly (but only through the endogenous variables $first\_flood$, $second\_flood$, $loss\_big\_ff$ and $loss\_big\_sf$) we will be able to reduce the location bias.

For this purpose we use variables that describe house location (slope steep or flat) and house characteristics (the presence of elevated floor or cellar) to instrument for variables $first\_flood$, $second\_flood$, $loss\_big\_ff$ and $loss\_big\_sf$. We estimate the probit model with endogenous covariates in equation (4.2) using the two-step estimator suggested by Newey (1987). The estimates are presented in panel three of Table 8. The signs of parameters do not change, but their significance drops. Wald chi-squared test of exogeneity does not reject the null hypothesis. Given this, the IV estimation is less efficient than OLS leading to inflated standard errors.

4.4.3 Commuting distance

In the previous section we estimated the extensive margin and found that exposure to floods affects the decision to commute in a nonlinear manner. In this section we research the intensive margin to learn whether affected individuals commute shorter or longer distances. We estimate the following model:

$$\begin{aligned} distance_i &= \gamma_0 + \gamma_1 first\_flood_i + \gamma_2 second\_flood_i + \\ &+ \gamma_3 loss\_big\_ff_i + \gamma_4 loss\_big\_sf_i + \gamma_5 cov\_more\_ff_i + \gamma_6 cov\_more\_sf_i + \\ &+ \gamma_7 educ2_i + \gamma_8 educ3_i + \gamma_9 male_i + \gamma_{10} age30_i + \gamma_{11} age40_i + \gamma_{12} age50_i + \\ &+ \gamma_{13} married_i + \gamma_{14} kids1_i + \gamma_{15} kids2_i + \gamma_{16} kids3_i + \\ &+ \gamma_{17} fam\_dist_i + \gamma_{18} \lambda_{dist,\,i} + \mu_i, \end{aligned} \tag{4.3}$$

where $distance$ is commuting distance in km, $\lambda = \frac{\phi(\cdot)}{\Phi(\cdot)}$ is the inverse Mill's ratio estimated from the selection equation (4.2). fam_dist is a variable that describes distance traveled by a family member who started commuting before the respondent. This variable is analogous to fam_com from the probit regression (4.2). All other variables are defined in Table D.2.1. As an exclusion restriction in the selection equation we use variable fam_com. The estimates of regression (4.3) are shown in Table 9.

Table 9: OLS estimates of equation (4.3).

Variable	Estimate		SE
first_flood	-0.082		(1.01)
second_flood	-3.624	**	(1.54)
loss_big_ff	1.076		(1.27)
loss_big_sf	0.532		(2.78)
cov_more_ff	-2.587		(1.88)
cov_more_sf	-2.826		(2.86)
married	0.908		(1.16)
male	1.678	**	(0.76)
age30	1.617		(1.67)
age40	2.230		(1.53)
age50	4.045	***	(1.44)
educ2	0.471		(0.89)
educ3	5.460	**	(2.35)
kids1	-1.043		(1.10)
kids2	-0.443		(1.86)
kids3	2.304		(1.40)
fam_dist	0.876	***	(0.07)
λ_{dist}	4.863	***	(1.07)
cons	-10.194	***	(3.26)

Notes: The regression includes region (obec) dummy variables. The number of observations is 114 and adj. R^2 = 0.55. Standard errors are clustered by $family\_id$. *** - 1%, ** - 5%, * - 10% significance levels.

The estimates suggest that individuals affected by one flood commute less than unaffected individuals, although this difference is not statistically significant. Individuals affected by two floods commuted on average 3.6 km less than unaffected commuters. Signs on variables that measure flood related losses and insurance settlements have the same signs as in Table 8.

Individuals with higher losses commute slightly longer distances and individuals with high settlements commute shorter distances. Males, educated and older individuals commute to farther cities. New commuters commute slightly shorter distances than more experienced members of their families. This supports the fact that two commuters from the same family often work in the same city or share the same car. Further, the significance of the inverse Mill's ration means that there is strong selection into commuting; the decision to start commuting and how far to commute are two interrelated decisions.

4.5 Conclusion

In this paper we find that the patterns of commuting for work among Czech households living in flood risk areas are affected by their exposure to floods. The effect is nonlinear: an average individual affected by one flood is 11.2% more likely to commute. Since commuting is on average associated with higher income, it allows affected individuals to accumulate resources to cope with flood related losses. Many respondents from affected families expressed unhappiness about living in risk areas, because their assets are often damaged by rising water from the River Bečva or flash floods. These respondents would like to acquire anti-flood adaptation measures to reduce their households' vulnerability to floods or move to safer areas. Commuting gives them such an opportunity and "successful" commuters do out-migrate eventually.

Those who stay commute less because they were not "good" commuters in the first place. Compared to unaffected individuals, respondents who have experienced at least two floods are 20.2% less likely to commute and they do so 3.6 km less on average. This implies that people who stayed after two floods are indeed different from those who experienced only one or no floods in some fundamental characteristics related to education, experience, family circumstances and individual migration costs.

It must be clarified that individuals commute more after the first flood and less after the second flood. This is established from the regression estimates. However, out-migration is a phenomenon that we could identify but could not quantify during the study. It happens somewhere in between the first and second floods or straight after the second flood. Our perception is that the out-migration from the surveyed area is a rational decision rather than a need to flee from flood devastated areas. None of the floods, except for the one in 1997, was devastating enough to generate such an effect.

Respondents with assets badly affected by floods are more likely to commute, though this difference is not significant. Residents who obtained generous insurance settlements are significantly less likely to commute. This suggests that commuting and insurance settlements are substitutes; they help affected households cope with flood losses and decrease vulnerability to flood risk.

We also find that the decision to commence commuting is to a large extent determined by the presence of a commuter in the family. An individual with an active commuter is 53.8% more likely to start commuting than somebody without a commuter. To decrease transportation costs commuters might share transportation means and thus choose to work in similar destinations. On average, a 1 km increase in distance commuted by a family member is associated with an 0.87 km increase in distance commuted by a new commuter within the same family.

We can think of several follow-up studies. The decision to commute or out-migrate depends on individual attitude to risk, whereby less risk averse individuals might be more likely to commute due to uncertainty. Also, it would be interesting to research housing prices and the characteristics of individuals who have permanently out-migrated from the surveyed risk areas. Comparing their characteristics with those of stayers will shed light on determinants of permanent out-migration. Further, researching insurance contracts in more detail should unveil a pattern as to whether partially settled claims were the result of individual negligence and under-insurance or companies' parsimony.

Bibliography

Barrios, S., Bertinelli, L., Strobl, E., 2006. Climatic change and rural-urban migration: the case of sub-Saharan Africa. Journal of Urban Economics 60 (3), 357–371.

Borga, M., Anagnostou, E., Bloschl, G., Creutin, J.-D., 2011. Flash flood forecasting, warning and risk management: the HYDRATE project. Environmental Science & Policy 14 (7), 834–844.

Botzen, W., Aerts, J., van den Bergh, J., 2009. Willingness of homeowners to mitigate climate risk through insurance. Ecological Economics 68 (8–9), 2265–2277.

Brázdil, R., Řezníčková, L., Valášek, H., Havlíček, M., Dobrovolný, P., Soukalová, E., Řehánek, T., Skokanová, H., 2011. Fluctuations of floods of the River Morava (Czech Republic) in the 1691–2009 period: interactions of natural and anthropogenic factors. Hydrological Sciences Journal 56 (3), 468–485.

CHMI, 2012. Projects of flood effects assessment (in Czech). Czech Hydrometeorological Institute, Prague.

Christensen, J. H., Christensen, O. B., 2003. Climate modelling: severe summertime flooding in Europe. Nature 421 (6925), 805–806.

Grothmann, T., Reusswig, F., 2006. People at risk of flooding: Why some residents take precautionary action while others do not. Natural Hazards 38 (1–2), 101–120.

Heckman, J. J., 1979. Sample selection bias as a specification error. Econometrica 47 (1), 153–161.

Kreibich, H., Müller, M., Thieken, A. H., Merz, B., 2007. Flood precaution of companies and their ability to cope with the flood in August 2002 in Saxony, Germany. Water Resources Research 43 (3).

Kreibich, H., Thieken, A., 2009. Coping with floods in the city of Dresden, Germany. Natural Hazards 51 (3), 423–436.

Kundzewicz, Z. W., Pińskwar, I., Brakenridge, G. R., 2013. Large floods in Europe, 1985–2009. Hydrological Sciences Journal 58 (1), 1–7.

Newey, W. K., 1987. Efficient estimation of limited dependent variable models with endogenous explanatory variables. Journal of Econometrics 36 (3), 231–250.

Reuveny, R., 2007. Climate change-induced migration and violent conflict. Political Geography 26 (6), 656–673.

Thieken, A. H., Kreibich, H., Müller, M., Merz, B., 2007. Coping with floods: preparedness, response and recovery of flood-affected residents in Germany in 2002. Hydrological Sciences Journal 52 (5), 1016–1037.

Wachinger, G., Renn, O., Begg, C., Kuhlicke, C., 2013. The risk perception paradox – implications for governance and communication of natural hazards. Risk Analysis 33 (6), 1049–1065.

Warner, K., 2010. Global environmental change and migration: governance challenges. Global Environmental Change 20 (3), 402–413.

Yeo, S., 2002. Flooding in Australia: A review of events in 1998. Natural Hazards 25 (2), 177–191.

Appendix

D.1 Map of surveyed area

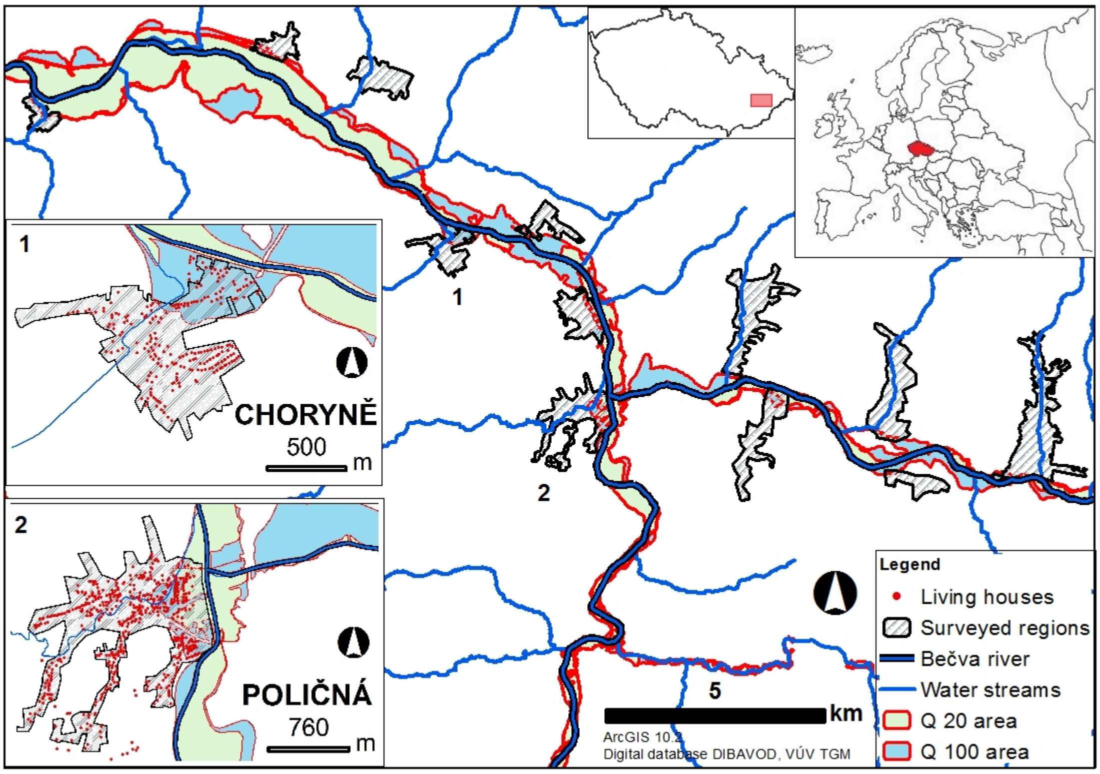


Figure D.1.1: Population of interest. Authors' illustration.

D.2 Definitions of variables

Table D.2.1: Definitions of covariates in regressions (4.1), (4.2) and (4.3).

Variable	Definition
commute	= 1 if a respondent started commuting after a respective flood date; and 0 otherwise.
first_flood	= 1 if a respondent experienced only one flood; and 0 otherwise.
second_flood	= 1 if a respondent experienced at least two floods; and 0 otherwise.
loss_big_ff	= 1 if total reported losses after the first flood exceed EUR 2k; and 0 otherwise.
loss_big_sf	= 1 if total reported losses after the second flood exceed EUR 2k; and 0 otherwise.
cov_more_ff	= 1 if the insurance company covered more that 50% of claimed losses after the first flood; and 0 otherwise.
cov_more_sf	= 1 if the insurance company covered more that 50% of claimed losses after the second flood; and 0 otherwise.
age	continuous variable that measures reported individual's age.
age2	$= \text{age}^2$.
age30	= 1 if respondent's age is in range (20 30]; and 0 otherwise.
age40	= 1 if respondent's age is in range (30 40]; and 0 otherwise.
age60	= 1 if respondent's age is in range (50 60]; and 0 otherwise.
exper	continuous variable that measures reported individual's work experience.
educ2	= 1 if an individual has complete secondary education or vocational training; and 0 otherwise.
educ3	= 1 if an individuals holds a Bachelor's degree or above; and 0 otherwise.
married	= 1 if the respondent is married; and 0 otherwise.
kids1	= 1 if there is one child in the family; and 0 otherwise.
kids2	= 1 if there are two children in the family; and 0 otherwise.
kids3	= 1 if there are three children in the family; and 0 otherwise.
male	= 1 if a respondent is male; and 0 otherwise.
occ_type2	= 1 if respondent's occupation is in the medium-skilled category; and 0 otherwise.
occ_type3	= 1 if respondent's occupation is high-skilled; and 0 otherwise.
occ_type4	= 1 if a respondent is an entrepreneur; and 0 otherwise.
fam_com	= 1 if respondent's family has another commuter who started commuting first; and 0 otherwise.
fam_dist	continuous variable that measures commuting distance (in km) for an individual who started commuting first.
λ_{report}, λ_{dist}	inverse Mill's ratios, $\lambda = \frac{\phi(\cdot)}{\Phi(\cdot)}$.