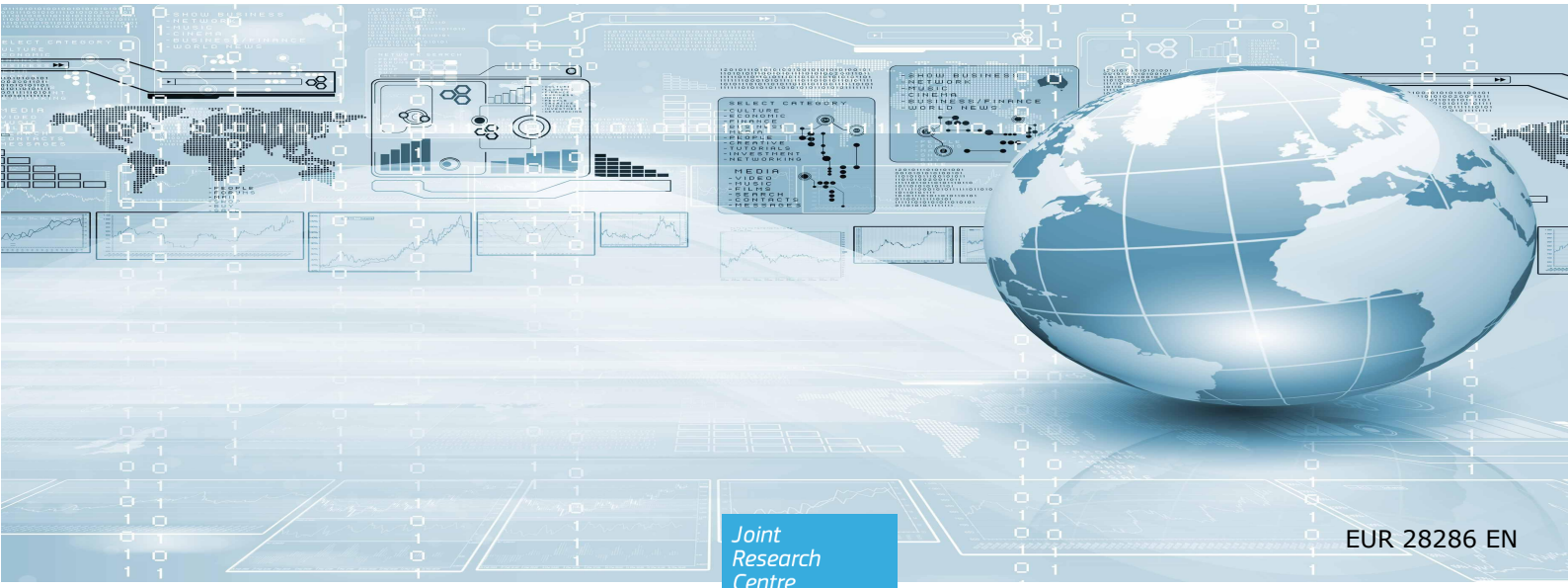
European Commission

JRC TECHNICAL REPORTS

Migration within the EU: the role of education, wage differences and cultural barriers

Damiaan Persyn

2017



Joint Research Centre

EUR 28286 EN

Contact Information
Name: Damiaan Persyn
Address: Joint Research Centre, Edificio Expo. c/ Inca Garcilaso, 3. E-41092 Seville, Spain
E-mail: damiaan.persyn@ec.europa.eu
Tel.: +34 9544-88232
JRC Science Hub
https://ec.europa.eu/jrc

JRC104494

EUR 28286 EN

PDF ISBN 978-92-79-64431-3 ISSN 1831-9424 doi:10.2791/373605

Luxembourg: Publications Office of the European Union, 2017

How to cite: Damiaan Persyn (2017). "Migration within the EU: the role of education, wage differences and cultural barriers", JRC Technical Report JRC104494, European Commission, DG Joint Research Centre, EUR 28286 EN, doi:10.2791/373605

Migration within the EU: the role of education, wage differences and cultural barriers

Abstract
There exist marked differences in the educational attainment of immigrants, depending on both the level and distribution of income in the country of origin and destination. This paper estimates an education-specific gravity equation for migration between European countries. Given the lack of data on migration flows by level of education, these are proxied by the difference in resident migrants by nationality and level of education, between the years 2000 and 1990. I find that highly educated individuals are more likely to migrate. They are less sensitive to geographical and cultural distance as barriers to migration, but are not unambiguously more responsive to wage differentials. Controlling for education-specific wage differences between origin and destination removes only part of the observed differences in migration behaviour between education groups.

Migration within the EU: the role of education, wage differences and cultural barriers

Damiaan Persyn

Contents

Abstract

There exist marked differences in the educational attainment of immigrants, depending on both the level and distribution of income in the country of origin and destination. This paper estimates an education-specific gravity equation for migration between European countries. Given the lack of data on migration flows by level of education, these are proxied by the difference in resident migrants by nationality and level of education, between the years 2000 and 1990. I find that highly educated individuals are more likely to migrate. They are less sensitive to geographical and cultural distance as barriers to migration, but are not un-ambiguously more responsive to wage differentials. Controlling for education-specific wage differences between origin and destination removes only part of the observed differences in migration behaviour between education groups.

JEL codes: F22 - J61 - O15 - C25

Keywords: International migration, Random utility model, Education

1. Introduction

This paper considers differences in migration behaviour between individuals with different levels of education, for the case of migration between developed countries. We use a gravity equation for migration based on a discrete choice framework, and approximate migration flows using data on the stock of migrants in EU15 member states, plus Iceland and Norway. This selection of countries is based on data availability, and the aim to estimate migration flows between countries where migration has been free over the period considered (1990-2000).[1]

People with a higher level of education have been observed to be more likely to migrate between US states (Molloy et al., 2011), and between EU regions (Mauro and Spilimbergo, 1999), even when controlling for a variety of characteristics, such as age, marital status or home ownership. An important question is whether educated people face different circumstances (for example, higher wage differentials between origin and destination), or whether this different behaviour even exists when controlling for these circumstances, and are due to, for example, residual differences in 'taste' for migration which correlate with the level of education. A simple example would be the case of log-linear utility, which would imply that (abstracting from migration costs) if the income difference between two regions is identical for low and highly educated individuals, this would imply a larger difference in utility terms for people with a lower level of education, due to the concavity of the utility function. Another example would be the presence of fixed costs of migration in case of linear utility, which would mean that individuals with high wages would be more prone to migrate given a certain *relative* wage differential, as for them, the fixed costs of migration would be smaller compared to the potential income gain. There may be numerous other considerations, such as financial constraints and capital market imperfections, which would imply that individuals with low incomes migrate less than expected even in the face of large interregional wage differentials.

To avoid imposing restrictions regarding the functional form of the effect of wages and wage differentials, I start with an exploratory data analysis, using semi-parametric methods to unveil the relation between wage differentials and migration. I subsequently impose more restrictions on the data in order to obtain a migration equation which is as general as possible, but simultaneously sufficiently succinct to allow to state and test clear hypothesis about differences between education groups regarding their migration behaviour. The obtained empirical results will be used to introduce differences in the migration behaviour by education level in the regional CGE model Rhomolo (Mercenier et al., 2016).

2. Data

There exists no comprehensive dataset on migration flows between EU member states. Although the EU micro-level Labour Force Survey is extensive, it is unsuited for analysing migration. As argued by Martí and Ródenas (2007), the only question asked in the LFS which in theory could be used to determine historical migration movements, asks all individuals about their residence exactly one year before the date of the interview. This is problematic for at least two reasons. First, in most countries the panel is rotating. Individuals being interviewed for the second year will therefore be found to be internationally immobile simply because of the sampling design. Second, in most countries sampling is based on population registries. The longer it takes between arriving in the country of destination, and before a migrant enters into the registry, is sampled, and finally interviewed, the less the respondents will appear to be internationally mobile. Realistically, there might be a gap of at least several months between arriving and being interviewed. Moreover, as both the exact procedures for registration and the LFS sampling method differ significanlty between countries, so will the magnitude of the resulting bias. These issues render the question about the whereabouts of the respondent one year before the interview useless for analysing international migration flows.

The EU statistical office Eurostat used to report administrative data on international and interregional (within-country) migration flows for a selection of member states, but has stopped to do so and has removed these datasets from its publicly available database.

[1]The countries we consider are the major EEC and EFTA members in 1990, between which there existed free movement of workers and/or people, to various degrees, but especially so from 1992/1993 onwards with the establishment of the EU and EEA. Our data on the number of residents with a foreign nationality stems from Docquier et al. (2007).

Although this administrative data would be a quite precise source to analyse regular migration (and was used in Brandsma et al., 2014, for example), it lacks information on the educational background information, and therefore is unsuitable for this study.

Unfortunately, there exists no data on bilateral flows by level of education. Therefore, to approximately calculate the migration flows between member states by education level, I have to resort to taking the difference in the stocks of migrants by nationality, between the years 2000 and 1990, as reported by Docquier et al. (2007). Although it might be the only route to obtain approximate migration flows by education level, the approach is problematic for several reasons. Firstly, it attributes all changes in stocks of migrants by nationality to migration, ignoring other demographic factors, such as mortality and fertility. This is problematic in the context of studying differences in migration behaviour between educational groups, since these demographic factors (and the bias in the estimation of migration flows) may differ significantly between them. It is reasonable to assume, for example, that the average age of the individuals with the lowest level of education is significantly higher compared to those with medium or high levels. If this is the case, changes in stocks of migrants with low levels of education may rather reflect higher mortality rates in this group, or issues such as return-migration, or naturalisation (when migrants obtain the nationality of the host country) which we will ignore in this study. Also from a behavioural point of view, when considering migration as an investment decision, it is to be expected that, ceteris paribus, older subgroups are less prone to migrate.

Table 1 shows the stock of immigrants by nationality in the years 2000 and 1990, for the subset of EEA countries contained in the dataset of Docquier et al. (2007). I chose the EEA as the relevant set of countries, as these are all countries between which there has been long-standing agreements assuring the free movement of labour. Switzerland was excluded since it agreed to free movement of labour from EU countries only in the year 1999. In some countries, the number of residents with a specific nationality decreased over time, sometimes considerably so. I will ignore these observations.

3. Methodology

To derive an estimable migration equation relating the aggregate inter-regional migration flows to behavioural parameters I use the approach described in Brandsma et al. (2014), and start from the individual migration decision. Indices for the level of education are left out for now. Consider an individual k from origin region o, maximising indirect utility, V_{kor}, across all possible destinations r. Destination d will be chosen if

$$V_{kod} > V_{kor} \quad \forall\, r$$
$$V_{kod} = Z_{od}\beta + \xi_{od} + e_{kod}.$$

The indirect utility V_{kod} of worker k migrating from origin region o to destination region d is determined by characteristics, Z_{od}, of regions o and d. The vector β contains the slope parameters and the vector product $Z_{od}\beta$ represents the utility the individual receives from the characteristics contained in the vector Z_{od}. The scalar error term, ξ_{od}, represents unobserved location and origin-destination-pair characteristics. $Z_{od}\beta$ and ξ_{od} assign the same utility level to all individuals considering migration from o and to d. The idiosyncratic error term e_{kod} varies across both individuals and regions. It accounts for the fact that not all individuals from the same origin region choose the same destination. The probability that location d is chosen as a destination by a resident k of region o, $\Pr(M_{kod} = 1)$, then equals

$$\Pr(M_{kod} = 1) = \Pr(V_{kod} > V_{kor}) \qquad \forall\, d \neq r$$
$$\Pr(M_{kod} = 1) = \Pr(e_{kod} - e_{kor} > Z_{or}\beta - Z_{od}\beta + \xi_{or} - \xi_{od}) \qquad \forall\, d \neq r.$$

Now assume that the idiosyncratic error term e_{kod} follows an iid extreme value distribution. McFadden (1973) shows this yields the following probability for a worker k to migrate from o to d:

$$\Pr(M_{kod} = 1) = \frac{\exp(Z_{od}\beta + \xi_{od})}{\sum_{d=1}^{R} \exp(Z_{od}\beta + \xi_{od})}. \tag{1}$$

Berry (1994) in turn shows that probability (1) of migrating from o to d can be interpreted as the share of workers from o migrating to d. Following Sorensen et al. (2007), we therefore

	AT	BE	DE	DK	ES	FI	FR	GR	IE	IS	IT	LU	NL	NO	PT	SE
AT		632	18000	-38	2123	84	-301	34	243	23	2480	107	350	91	123	-590
BE	872		3501	100	12344	62	-19	30	383	24	2114	3362	301	137	527	190
DE	73956	8223		325	59654	1158	9502	274	3071	42	13188	1102	-8937	2450	3574	-440
DK	612	751	5524		2712	146	1076	17	229	213	603	314	679	2430	74	-5190
ES	-1331	-3459	25613	547		294	-75356	19	2758	22	4659	252	2524	283	1634	200
FI	671	1520	5256	67	2442		784	18	465	96	512	362	609	2087	94	-17060
FR	3131	2723	34275	889	72112	456		102	3119	5	10937	4764	3092	582	5593	1330
GR	1826	-283	44000	156	450	183	-1171		150	31	2148	231	1327	114	35	-1830
IE	305	769	5165	204	1501	113	1104	12	0		558	-42	324	67	176	430
IS	62	80	446	1423	84	53	90	1	36		14	16	85	1021	13	-260
IT	18421	-8624	106000	662	11301	452	-101037	105	1962	156		-4165	3927	336	641	70
LU	168	-746	936	10	404	6	-387	1	10	1	155		58	9	93	0
NL	3533	17467	9794	1373	0	290	5767	60	1220	1499	2334	356		912	1071	650
NO	-4179	351	1920	716	2490	100	196	10	187	91	273	2	356		90	-8490
PT	-2672	5901	46131	143	25059	59	26866	6	373	181	1139	11356	1544	144		30
SE	9	1108	6472	278	3504	8919	1448	34	549	295	931	435	758	9569	126	
UK	3027	1441	27000	2658	47953	1060	21751	276	80186	137	7599	827	5207	1070	1924	3340

Table 1: The change in the number of resident migrants (stock), between 1990 and 2000, by nationality (rows) and country of residence (columns). There is no data for the UK as a country of residence. *Source: Docquier et al. (2007)*

write the share of migrants from o to d as:

$$s_{od} = \Pr(M_{kod} = 1) = \frac{\exp(Z_{od}\beta + \xi_{od})}{\sum_{d=1}^{R} \exp(Z_{od}\beta + \xi_{od})}, \tag{2}$$

and the share of stayers in region o as:

$$s_{oo} = \Pr(M_{koo} = 1) = \frac{\exp(Z_{oo}\beta + \xi_{oo})}{\sum_{d=1}^{R} \exp(Z_{od}\beta + \xi_{od})}. \tag{3}$$

Dividing equation (2) by (3) and applying a logarithmic transformation yields a simple estimable migration equation:

$$\ln\left(\frac{s_{od}}{s_{oo}}\right) = \ln\left(\frac{\exp(Z_{od}\beta + \xi_{od})}{\exp(Z_{oo}\beta + \xi_{oo})}\right) = Z_{od}\beta - Z_{oo}\beta + \xi_{od} - \xi_{oo}. \tag{4}$$

In this paper, I do not consider nested choices as in Brandsma et al. (2014), as the origin-destination pairs are countries and have no clear nesting structure. The assumption of independence of irrelevant alternatives is therefore more likely to hold than in the case of regions nested in countries, which was considered in Brandsma et al. (2014).

4. The relationship between income differences and migration flows: exploratory analysis

This section considers the relationship between the income difference between two locations and the historic migration flow (as proxied by the difference in stocks) between them, by means of semi-parametric methods. These methods impose a standard rigid parametric structure on a subset of variables in the migration equation, while allowing the data to freely determine the shape of the relationship for some other variables. The econometric estimation equation derives directly from Eq. (4), but distinguishes between education levels $e \in \{Lo, Med, Hi\}$:

$$\begin{aligned}\ln\left(\frac{s_{ode}}{s_{ooe}}\right) = \beta_0 &+ \beta^{\text{dist}}\text{lndist}_{od} + \beta^{\text{clang}}\text{clang}_{od} + \beta^{\text{cbord}}\text{cbord}_{od} \\ &+ \sum_{r\in\{o,d\}} \beta_r^{\text{pop}}\text{lnpop}_r + \beta_r^{\text{tourism}}\text{tourism}_r + \beta_r^{\text{lnpli}}\text{lnpli}_r \\ &+ m(\text{wage}_{de} - \text{wage}_{oe}) + \xi'_{od}\end{aligned} \tag{5}$$

where $m()$ is a non-parametric function to be estimated. The variables considered are the geographic distance between the capitals (lndist), and two proxies for cultural distance: a dummy indicating a common language between origin and destination (clang), and a dummy indicating a common border (cbord). Furthermore, we include the population in origin and destination as a measure of size (lnpop); the number of overnight stays by tourists, per capita, in origin and destination, as a proxy of amenities such as climate (tourism); and the purchasing power price index to control for the cost of living (lnpli). Finally, we consider the average wage by education group ($\text{wage}_{e,r}$). By considering education-specific wages, rather than the economy-wide average wage, the analysis adresses some of the selection issues which are inherent in migration, albeit quite partially. Migrants are likely a non-random sample from the underlying population, for which the average wage in origin and destination is not the relevant reference wage.

Table 2 shows the results for the parametric part of the estimation. The specification in column (I) corresponds to equation 5 and includes $m(wage_{d,e} - wage_{o,e})$ nonparametrically. Column (II) rather uses $m(wage_{d,e})$ for the nonparametric part, and estimates a slope parameter for $wage_{o,e}$. The specification in column (III) considers the reverse: including $m(wage_{o,e})$ nonparametrically and including a linear effect for $wage_{d,e}$. For comparison, column (IV) is a traditionally specification, fully parametric, including linear effects for all variables. Summarising the effects of the parameter estimates over all specification, we find the expected effects on the variables proxying distance (also for common language and border which proxy cultural distance). The size variables (lnpop) have the expected sign, as do the variables proxying amenities (tourism_r). The price indexes also have the expected sign and are significant.

	(I)	(II)	(III)	(IV)
logdist_{od}	-0.877^{***}	-0.646^{***}	-0.589^{***}	-0.606^{***}
	(−5.05)	(−4.09)	(−3.84)	(−4.06)
clang_{od}	0.487	0.448	0.403	0.461^{+}
	(1.26)	(1.27)	(1.19)	(1.42)
cborder_{od}	1.016^{***}	1.071^{***}	1.098^{***}	1.139^{***}
	(3.43)	(3.89)	(4.14)	(4.68)
lnpop_{o}	-0.386^{***}	-0.422^{***}	-0.431^{***}	-0.415^{***}
	(−5.90)	(−5.36)	(−7.52)	(−6.25)
lnpop_{d}	0.725^{***}	0.715^{***}	0.723^{***}	0.717^{***}
	(10.88)	(12.36)	(7.50)	(11.64)
tourism_{o}	-0.159^{***}	-0.122^{***}	-0.162^{***}	-0.162^{***}
	(−3.75)	(−2.59)	(−4.38)	(−4.76)
tourism_{d}	0.145^{***}	0.156^{***}	0.104^{**}	0.152^{***}
	(3.87)	(4.68)	(2.20)	(5.98)
lnpli_{d}	-2.075^{**}	-3.617^{***}	-3.196^{***}	-3.841^{***}
	(−2.53)	(−4.68)	(−3.30)	(−4.89)
lnpli_{o}	1.118^{+}	−0.717	−0.797	−0.700
	(1.46)	(−0.83)	(−1.12)	(−0.96)
wage_{de}		0.185^{***}		0.198^{***}
		(7.07)		(7.93)
wage_{oe}			-0.0361^{+}	-0.0392^{+}
			(−1.47)	(−1.64)
_cons				9.130^{+}
				(1.48)
N	416	416	416	416
adj. R^2	0.526	0.568	0.476	0.580

t statistics in parentheses^{+} $p < 0.2$, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: A semiparametric gravity equation for migration: esimates of the slope parameters. Column (I) includes $m(wage_{d,e} - wage_{o,e})$ nonparametrically. Column (II) uses $m(wage_{d,e})$ and estimates a slope parameter for $wage_{o,e}$, and vice-versa in column (III). Column (IV) is fully parametric, including linear effects for all variables. Figure 1 shows the corresponding estimated nonparameteric part.

For the three estimates corresponding to the columns (I) to (III) of Table 2, Figure 1 shows the remaining variation in migration with the parametric part partialled out, overlaid with the estimated non-parametric function m(). The left panel corresponds to equation Eq. (5), where the difference in wages $m(\text{wage}_d - \text{wage}_o)$ enters nonparametrically. There is a clear upward slope: controlling for a large set of covariates, there is a clear positive relation between the wage difference between destination and origin, and migration. This effect seems broadly linear. The middle panel of the figure considers a specification with $m(\text{wage}_d)$ as the nonlinear part, and includes wage_o linearly. Again there is a clear positive relationship, which seems approximately linear. The right panel considers a nonlinear effect for the wage in the country of origin, through specifying $m(\text{wage}_o)$. Here, the slope is negative, but quite flat. The effect seems stronger for the lowest values of wage_o.

Visual inspection of the graphs reported in Figure 1 suggest the effect of wages on migration in equation 5 is approximately linear. Comparing the specification reported in column (IV) of table 2 with an alternative with wages entering in logarithms resulted in an adjusted R^2 of 0.567, which suggests a significantly worse fit as compared to the purely linear model with an R^2 of 0.58.

It is important to note that this exploratory analysis has ignored any differences between education groups regarding the effect of the explanatory variables on migration: although we consider wages which are specific to each education group, the slope parameters are identical across education groups. One possible explanation for the apparent weak relation between the origin-wage and migration (right panel of figure 1) therefore could be that there are underlying significantly different effects (slopes) for separate education groups.

Figure 1: The estimated non-parametric function $m()$, for three specifications of the semiparametric kernel regression Eq. (5). Dependent variable: $\ln(s_{ode}/s_{ooe})$. The left panel plots $m(wage_{d,e} - wage_{o,e})$. The middle panel corresponds to a specification with $m(wage_{d,e})$ and includes $wage_{o,e}$ linearly in the parametric part. The right panel shows $m(wage_{o,e})$ en includes $wage_{d,e}$ linearly.

Differentiating the effects per education group will be the topic of the next section.

5. Differentiating by level of education

5.1 Differentiating by level of education: Econometric analysis

The limited number of observations implies that we cannot hope to distinguish separate effects on all variables for all education groups. I therefore chose to restrict a common effect between education groups for a large subset of variables, while allowing only differences between education groups regarding the effect of wages in both origin and destination, and the effect of the contiguity dummy, as a proxy for cultural distance. The three variables proxying for migration costs: geographic distance, and the common language and the contiguity dummies, are too highly correlated to separately estimate while distinguishing between education groups. Given that the exploratory analysis suggests that the effect of income differences is approximately linear, I start by considering only linear effects when considering differences between education groups.

Table 3 shows the results. The first column shows a specification without education-specific variables (such as wages in origin and destination), but allows for a separate level effect per education group. The specification includes a constant, and leaves out the effect for the lowest level of education. The coefficients on the dummies for Medium and High levels of education (I(Med) and I(Hi) respectively) therefore have to be interpreted as expressing the difference of the level for these group relative to the base category of Low education. The high and significant coefficients on the Medium and High education dummy variable indicate that on average, controlling for the basic geography and distribution of the population, the medium and high educated groups are overrepresented in the migrant population, relative to the low educated. More specifically, $\ln(s_{ode}/s_{ooe})$ is estimated to be 0.75 higher for medium educated, compared to low educated. This implies that the ratio of shares of migrants versus non-migrants is exp(0.75), or rougly 2 times higher. Given that the share of stayers s_{ooe} is quite close to 1 for all education groups, across all origins, this equates to an estimated double share of migrants in the medium educated group compared to the low-educated. For the high-educated, this becomes $\exp(1.597) \approx 5$ times more. These are quite considerable differences.

Column (II) in table 2 also allows for a different level effect depending on the education group, but now controls for education specific wages in origin and destination as explanatory factors of migration. The signs are as expected. Notably, when comparing to column (IV) in table 2 which is equal apart from the absence of different level effects by education, the effect of wages in the origin becomes stronger after including education-group-dummies. Moreover, the results suggest that the higher tendency to migrate which is observed for the highly educated, cannot be explained by the average wage differences between countries. To the contrary, controlling for wages in origin and destination, the estimation results predict predict higher shares of migrants among the medium and highly educated, as the estimated level effects for these groups become larger. This is a perhaps suprising result, as it runs counter to the intuition that higher average international wage differences (in absolute numbers) for highly educated can explain the relative overrepresentation of this group among the migrant population.

The specification in column (III) allows for differences in the overall level of migration

	(I)		(II)		(III)		(IV)	
$\ln pop_o$	−0.456***	(−6.81)	−0.415***	(−6.56)	−0.423***	(−6.67)	−0.387***	(−6.31)
$\ln pop_d$	0.769***	(12.31)	0.727***	(11.96)	0.732***	(12.17)	0.717***	(11.63)
$logdist_{od}$	−0.848***	(−5.98)	−0.833***	(−5.45)	−0.848***	(−5.69)	−0.860***	(−5.92)
$clang_{od}$	0.462^{+}	(1.40)	0.475^{+}	(1.49)	0.463^{+}	(1.51)	0.475^{+}	(1.53)
$cborder_{od}$	1.082***	(4.38)	1.012***	(4.29)	1.335***	(4.22)	1.304***	(4.04)
$cbord_{o,d}$ × I(Med)					−0.197	(−0.51)	−0.152	(−0.38)
$cbord_{o,d}$ × I(Hi)					−0.742*	(−1.84)	−0.693*	(−1.70)
$tourism_o$	−0.152***	(−4.46)	−0.165***	(−4.90)	−0.164***	(−4.85)	−0.148***	(−4.45)
$tourism_d$	0.139***	(5.12)	0.147***	(5.67)	0.142***	(5.56)	0.138***	(5.18)
$\ln pli_o$	−1.777***	(−2.77)	1.097^{+}	(1.43)	1.186^{+}	(1.53)	1.492*	(1.83)
$\ln pli_d$	0.322	(0.58)	−1.671*	(−1.78)	−1.137	(−1.20)	-1.431^{+}	(−1.36)
$wage_{o,e}$			−0.134***	(−4.77)	−0.104**	(−2.12)		
$wage_{o,e}$ × I(Med)					−0.103*	(−1.81)		
$wage_{o,e}$ × I(Hi)					0.00146	(0.03)		
$wage_{d,e}$			0.0824**	(2.15)	−0.0186	(−0.32)		
$wage_{d,e}$ × I(Med)					0.105**	(2.02)		
$wage_{d,e}$ × I(Hi)					0.117**	(2.48)		
$\ln wage_{o,e}$							−1.023*	(−1.90)
$\ln wage_{o,e}$ × I(Med)							−1.455**	(−2.35)
$\ln wage_{o,e}$ × I(Hi)							−0.767	(−1.25)
$\ln wage_{d,e}$							0.0569	(0.12)
$\ln wage_{d,e}$ × I(Med)							0.993**	(1.99)
$\ln wage_{d,e}$ × I(Hi)							1.668***	(3.08)
constant	−2.135	(−0.35)	−5.728	(−0.85)	−7.718	(−1.16)	−7.160	(−1.07)
I(Med)	0.750***	(4.27)	0.850***	(4.35)	1.001	(0.97)	2.109	(0.98)
I(Hi)	1.597***	(9.69)	1.974***	(4.72)	0.414	(0.40)	−0.397	(−0.16)
adj. R^2	0.581		0.603		0.616		0.616	

t statistics in parentheses. $^{+}$ $p < 0.2$, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: A gravity equation for migration, analysing the effect of education. Dependent variable: $\ln(s_{ode}/s_{ooe})$. (I): level effect for education, no wages. (II): education-specific wages and level effects for education. (III): allowing the effect of wages and the common border dummy (cborder) to differ between education groups. (IV): wages in logs.

depending on the level of education, as before, but additionally allows the effect of wages (both in origin and destination) and the common border dummy (cborder) to differ by the level of education. Again, low-education has been chosen as the base category, and the results of the interaction variables have to be interpreted as differences relative to this group. The reported effects of the variables reflect the values for the low-education group. For example, considering column (III) and approximating the share of stayers s_{ooe} by 1, the presence of a common border is expected to increase the share of low educated migrants s_{odLow} by $exp(1.335) \approx 3.8$. This increase is predicted to be $exp(1.335 - 0.197) \approx 3.1$ for the medium educated, and $exp(1.335 - 0.742) \approx 1.8$ for the highly educated, but only the latter difference (between high and low educated) is significant on the 5 percent level. Considering countries at equal distances (which is the relevant case, since the specification controls for geographic distance), I interpret the presence/absence of acommon border as a proxy for cultural proximity/distance of both countries. The results suggest that these factors matter less for more highly educated individuals.

The specification of column (IV) and (III) is identical, apart from the fact that wages now enter log-linearly. Given the identical value of the adjusted R^2-statistic, the quality of fit turns out to be quite similar to the linear case. This may be due to the fact that for some education groups, logs are the better fit compared to a linear specification, but the opposite holds vice-versa for other education groups.

5.2 Differentiating by level of education: Discussion

What pertains from the econometric analysis in the previous section is that wages in the origin are about equally important to the low and highly educated, but appear more important to the medium educated. When considering relative differences in wages (specification IV), there is some weak evidence that also highly educated value wage differences more compared to the low educated, but not as much as the medium educated.

Regarding the wage in the destination, we find that it does not have a significant effect on the low-educated, but more so (and about equally), for medium and highly educated. In log-terms, the effect is significantly higher for the highly educated.

It is important to reflect on what may be the underlying cause of these effects. One possible explanation for the lack of responsiveness of the lowest education levels to wage differences, would be financial constraints: although local wages are low and wages abroad are high, people lack the resources to migrate. As financial markets are imperfect, people can not always borrow against a possible future income stream from migration, and therefore the migration investment never takes place.

A more mundane explanation, however, lies with the limitations of using changes in the stocks of foreign nationals between 2000 and 1990 to approximate migration, our dependent variable in the econometric analysis. It is very likely that the group with the lowest education levels also has a substantial higher average age (since the average level of education has been steadily increasing over time). If this is the case, the low educated individuals are also older, and less likely to migrate because of several reasons (it is well established theoretically and empirically that migration is less likely beyond a certain age). Apart from age, there are other covariates which one would like to control for as they might affect migration decisions and are possibly correlated with the level of education, such as home ownership. Unfortunately, we cannot estimate or control for these effect as we do not have the required microeconomic background information.

Regarding the lower sensitivity of the highly educated to the wage in the region of origin, when compared to medium educated, there are likewise a number of possible explanations. A straightforward explanation would be concave utility: at higher levels of wages (and higher educated have higher average wages), the marginal utility of income decreases, and therefore the locally estimated slope of the linear relationship between local wages and migration is lower for higher educated as compared to medium educated: for medium educated, a 1 EUR decrease in the local wage represents a larger relative amount, compared to the highly educated. But the lower effect for high-educated in response to changes in the wages in the origin remains also when specifying wages in logs (specification IV).

Also interesting are the effect on the distance measures. Because these variables are correlated, it is impossible to differentiate effect per education group for all of them. We chose to report separate effects only for the common border dummy. The absence of a common border (the effect of which is the opposite of the reported effect), has a smaller negative effect on migration for the group of highly educated individuals.

6. Conclusion

The analysis of migration by level of education within the EU is greatly hampered by the lack of appropriate data. Considering the fact that migration flows are approximated by considering changes in stocks of migrants, care has to be taken when drawing conclusions from this study. Keeping this in mind, what pertains from the results, is that the effect of wages on migration appears approximately linear. In the analysis which did not allow for behavioural differences between education groups (table 2), the effect of the wage in the destination seemed stronger compared to the effect of the wage in the origin. When differentiating between education groups, individuals with low levels of education appear quite insensitive to wage differentials in the possible destinations. They are much more sensitive to local conditions, in contrast. When comparing across education groups, the medium-educated individuals appear most sensitive to wage differences in the origin. But more education always seems to imply a larger sensitivity to destination wages.

The results also suggest that language differences act as a hurdle to migration, although language may well proxy other cultural factors which could cause individuals to prefer more similar destinations, as compared to destination countries which are culturally more distant. Sharing a common border also appears to increase migration flows between countries. For the highly educated, having a common border matters significantly less for migration.

References

Berry, S. T., 'Estimating discrete-choice models of product differentiation', *The RAND Journal of Economics*, 1994, pp. 242–262.

Brandsma, A., Kancs, d. and Persyn, D., 'Modelling migration and regional labour markets: an application of the new economic geography model Rhomolo', *Journal of Economic Integration*, Vol. 29, No 2, 2014, pp. 372–406.

Docquier, F., Lowell, B. L. and Marfouk, A., 'A gendered assessment of the brain drain', IZA Discussion Papers 3235, Institute for the Study of Labor (IZA), 2007.

Martí, M. and Ródenas, C., 'Migration estimation based on the labour force survey: An EU-15 perspective', *The International Migration Review*, Vol. 41, No 1, 2007, pp. 101–126.

Mauro, P. and Spilimbergo, A., 'How do the skilled and the unskilled respond to regional shocks? the case of spain', *IMF Staff Papers*, Vol. 46, No 1, 1999, pp. 1–17.

McFadden, D., 'Conditional logit analysis of qualitative choice behavior'. In 'Frontiers in Econometrics', , edited by P. ZarembkaNew York: Academic Press, 1973.

Mercenier, J., Brandsma, A., Di Comite, F., Diukanova, O., Lecca, P., Lopez-Cobo, M., Monfort, P., Persyn, D., Rillaers, A., Thissen, M. and Torfs, W., 'Rhomolo-v2 model description: A spatial computable general equilibrium model for EU regions and sectors', JRC Technical reports JRC100011, Joint Research Centre (Seville site), 2016.

Molloy, R., Smith, C. L. and Wozniak, A., 'Internal migration in the United States', *The Journal of Economic Perspectives*, Vol. 25, No 3, 2011, pp. 173–196.

Sorensen, T., Fishback, P., Allen, S. and Kantor, S., 'Migration creation, diversion, and retention: New deal grants and migration 1935-1940', Nber working paper series, National Bureau of Economic Research, 2007.

List of figures

List of tables

Europe Direct is a service to help you find answers to your questions about the European Union.

Freephone number (*)**:**

00 800 6 7 8 9 10 11

(*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

More information on the European Union is available on the internet (http://europa.eu).

HOW TO OBTAIN EU PUBLICATIONS

Free publications:

- one copy:
  via EU Bookshop (http://bookshop.europa.eu);
- more than one copy or posters/maps:
  from the European Union's representations (http://ec.europa.eu/represent_en.htm);
  from the delegations in non-EU countries (http://eeas.europa.eu/delegations/index_en.htm);
  by contacting the Europe Direct service (http://europa.eu/europedirect/index_en.htm) or
  calling 00 800 6 7 8 9 10 11 (freephone number from anywhere in the EU) (*).

  (*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

Priced publications:

- via EU Bookshop (http://bookshop.europa.eu).

LF-NA-28286-EN-N

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

Publications Office

doi:10.2791/373605

ISBN 978-92-79-64431-3