Chapter 5

EU Enlargement: Migration from New EU Member Countries[♣]

by

Mariola Pytlikova

CIM and Dept. of Economics
Aarhus School of Business, Prismet,
Silkeborgvej 2, 8000 Aarhus C, Denmark,
Phone: +45 89 48 61 79, Email: marp@asb.dk

January 2006

Abstract: The main purpose of the paper is to give predictions of the migration potential from the 7 new EU member countries to the EEA/EU-13 countries. Being able to analyze "real" migration behavior from these particular countries over the period 1990-2000 helps me to avoid problems related to (double) out-of sample forecasts and to the assumption of invariance of migration behavior across a space that previous studies had to hold. Results of the econometric analyses reveal the importance of controlling for pairs of countries unobserved heterogeneity. Preliminary results regarding the predictions of future gross and net migration flows suggest that fears concerned with the large scale migration are hard to validate. Furthermore, current and predicted development in gross and net migration flows indicates that migration from the new EU member countries towards "old" EEA/EU countries is rather a temporary phenomenon in the sense that besides the inflows there will be substantial return flows as well.

Keywords: International migration, panel data, EU enlargement
JEL-code: J61, F22, O15

[♣] I would like to thank Jan Bentzen, Tor Eriksson, Peder J. Pedersen, Nina Smith and Valdemar Smith for fruitful discussions, Sebastian Buhai and participants at the seminar at the Department of Economics, Aarhus School of Business, for several helpful comments. Financial support from the Danish Social Science Research Council (Forskningsstyrelsen) is gratefully acknowledged.

5.1. INTRODUCTION AND MOTIVATION

On May 1st 2004, the European Union (EU) went through a process of the biggest enlargement in its history, when ten Central and Eastern European countries[73] joined the European Union.

In connection with the EU enlargement, labor migration issues have gained a special attention among the public, politicians and academics, as one of the fundamental rights guaranteed to EU citizens is free movement of workers between the member states. Given the geographical and cultural proximity and huge economic disparities, the old EU-15 countries have feared a mass migration from the new member states and its impact on their labor markets. In many of the “old” EU-15 countries, the likelihood of increased migration from the new members became an important issue for the internal political agenda. Consequently, a majority of the “old” member states have imposed restrictions on their employment and welfare systems driven by those fears of large-scale labor migration from Central and Eastern Europe (CEE). The only countries that have opened their labor markets to the 10 new EU members from the date of the EU enlargement in May 2004 are Sweden, the UK and Ireland. The rest of the EU countries decided to regulate the entry of the new members to their labor markets for at least two years. In 2006, the “old” EU-15 member states have an option to decide whether to extend the “transition period” for another three years. In principle, the “transition period” should end five years after the 2004 enlargement, but it may be prolonged for additional two years in the EU member states, where “migration might threaten to cause serious disturbances on the labor market” (European Commission, 2001). All in all, the “old” EU countries can keep their labor markets restricted to the new members up to 7 years from the enlargement.[74]

Given the sensitivity of this issue in the “old” EU member states there has been a strong need for analyses of the migration potential from Central and Eastern European countries (CEECs). Consequently, there is relatively rich empirical evidence focusing on identifying migration determinants and on forecasting the possible emigration pressure from CEECs. As regards the last one, forecasting of

[73] The following 10 countries have entered the EU on May 1st 2004: Cyprus, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, the Slovak Republic and Slovenia.
[74] There are, however, no such restrictions for Cyprus and Malta.

migration potential, there are, in general, two different approaches in the literature: surveys and econometric analyses. Surveys give estimates of CEE migration potential ranging from 6 to 30% of the populations, see e.g. Wallace (1998), Fassmann and Hintermann (1997). These numbers are clearly overestimated as only a minority of those who express an interest in migration actually migrates; see Fassmann and Hintermann (1997). Econometric analyses constitute the richest source of studies on this issue and their forecasts of CEE migration potential vary due to different modeling frameworks, estimation techniques or/and data samples. But the majority of existing studies forecast a long-run migration potential at around 3-4% of the source countries population. Taking into account out- and return migration, the net migration potential is usually estimated around 2% of source countries population, see Appendix Table 5.IA for an overview of existing studies and their forecasts, and further e.g. Dustmann et al. (2003) or Alvarez-Plata et al. (2003) for a more detailed literature review.

There are, however, several problematic issues connected to those studies and their estimates. Due to data limitations, the econometric analyses relied solely on out-of-sample historical data on migration[75] and/or past enlargement experience, and the estimates were further extrapolated to predict East-West migration, e.g. Bauer and Zimmermann, (1999), Boeri and Brücker (2001), Fertig (2001), Sinn et al. (2001), Alvarez-Plata et al. (2003), Zaiceva (2004). Thus, those studies assume the same migration behavior across the different countries, i.e. that migration decisions in the Central and Eastern European countries will respond to the same factors and in the same way as migration decisions in other source countries, e.g. Southern Europe or non-European countries like India and Pakistan with a very different economic and cultural background. Next, the studies assume invariance across time, i.e. that future migrations react to changes in economic factors in the same way as past migrations. Such assumptions of invariance across space and time are especially open to the Lucas critique. Some studies try to avoid the problems, at least partly, by controlling for unobservable country-specific effects. But the country-specific effects can not be used in out-of-sample predictions, therefore some studies try to model the differences in country-specific effects by time-constant factors that are characteristic

[75] I.e. migration waves from other countries than the CEECs.

for those countries or pairs of countries[76], see e.g. Fertig (2001), Dustman et al. (2003). Nevertheless, these variables can reveal a part of the unexplained variation between the countries or pairs of countries, but cannot get rid of it completely. Therefore they still suffer from an omitted variable bias. Consequently, the forecasts based on such (double) out-of-sample estimates might be seriously biased and do not clearly remove uncertainty connected to the expected migration flows from those countries. For a more detailed discussion of the problems with assessing the migration potential, see Dustmann et al. (2003) and Fertig and Schmidt (2000).

This paper brings a new aspect into the discussion as it presents analyses of East-West migration potential by using new sources of data. The dataset records migration patterns from the new EU member countries to the EEA/EU-13 countries over the period 1990-2000. To the best of my knowledge, this paper is the first that analyses East-West migration potential based on the actual Central and Eastern European migration behavior. This helps me to avoid the problems discussed above related to out-of-sample forecasts and the assumption of invariance of migration behavior across a space. Specifically, the data allows me to include the unobserved effects between the destination and particular Central and Eastern European countries into predictions of migration potential from those countries. Nevertheless, I still keep the assumption of invariance across time, i.e. that the migration behavior will not change with the openings of the EEA/EU-13 labor markets.[77]

Notably, the analyses reveal that unobservable effects of particular pairs of countries indeed play an important role in explaining emigration from those countries. The results of some preliminary predictions of future gross and net migration flows show that the net migration potential from the 7 new EU member countries is lower than the previous studies have estimated, with average annual net increases in stocks between 20,000 and 46,000 depending on the particular growth scenario used for the predictions. This leads to a total number of between 1.1 and 1.4 million migrants from those countries residing in 13 EEA/EU countries in 2015, which is equivalent

[76] Fertig (2001) followed by Boeri and Brücker (2001), Alvarez-Plata et al. (2003) and Zaiceva (2004) estimates fixed effects in the two-step procedure, where the unobservable country-specific effects are regressed by distance, language and development index (or other time-constant variables, e.g. traveling time, neighboring country etc. in further studies). Nevertheless, the authors are able to explain only 40-50% of country-specific effects.

[77] This is again open to the Lucas critique. The challenge to model and simulate the openings of the EEA/EU labor markets is left for the near future.

to 1.5–2% of the source countries' populations. On the other hand, the magnitude of the estimated gross migration flows is relatively high compared to the forecasts from previous studies. Specifically, the estimations show that the total accumulated gross flow over the period 2004-2015 is around 5.4–5.8% of the source population. Such a development in gross and net migration flows indicates that migration from the new EU member countries towards the "old" EEA/EU countries mostly is a temporary phenomenon.

The paper is organized as follows: Section 5.2 gives a description of migration trends from CEECs during the observed period 1990-2000. Next, in Section 5.3 a theoretical model of migration is presented. Section 5.4 describes the data used and presents econometric analyses and a discussion of estimation results. Section 5.5 gives some preliminary results of predictions of future gross and net migration flows from the new EU member countries. Finally, Section 5.6 concludes.

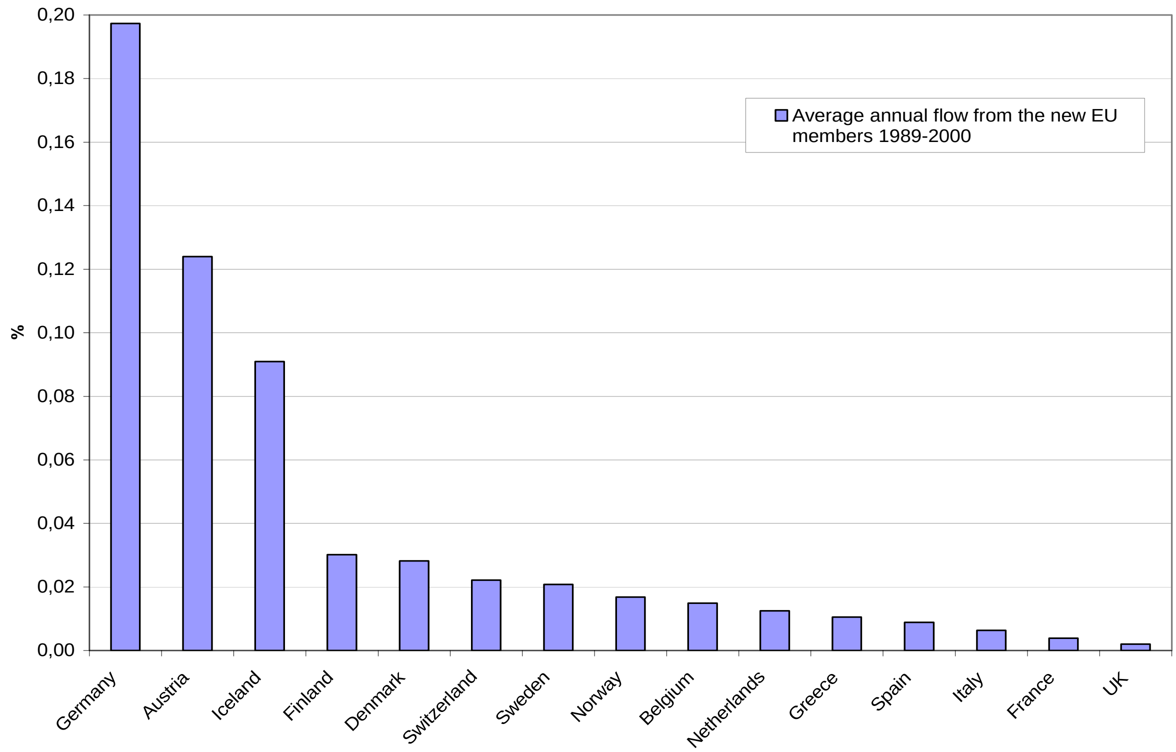
5.2. DEVELOPMENT OF MIGRATION FLOWS AND STOCK FROM CEE COUNTRIES AFTER 1989

Under the socialist regime, migration from the CEECs was tightly restricted and most of those who emigrated did so as political refugees. With the fall of the Iron Curtain in 1989, this situation changed and Central and Eastern Europeans became relatively free to migrate to other countries.[78] Many indeed have chosen to experience the newly acquired freedom of movement in order to improve their economic conditions or simply to experience living and working in another country without a fear of not being able to return and not to see relatives in their home countries again.

According to Figure 5.1, Germany and Austria not surprisingly experienced the highest average annual flows from the new EU member states both in absolute numbers and as a percentage of the particular host country population. The average annual inflow of immigrants from new EU member states over the period 1989-2000 was around 155,000 and 10,000, or 0.20 and 0.12% of the German and Austrian populations, respectively.

[78] Although "degrees of freedom" and "timings of freedom" were different across those countries.

Figure 5.1: *Immigration flows from the 7 new EU member states to the "old" EEA/EU countries as a percentage of the host countries' populations. Annual average 1989-2000.*



Note 1: Due to data availability the figure shows average migration flows to Italy from Poland only, to Spain from the Czech and Slovak Republics and Poland only. For the UK the figures show numbers of applicants for settlement and thus the gross migration flow is highly underrepresented. Ireland and Portugal have been excluded due to missing information on CEE countries of origin.

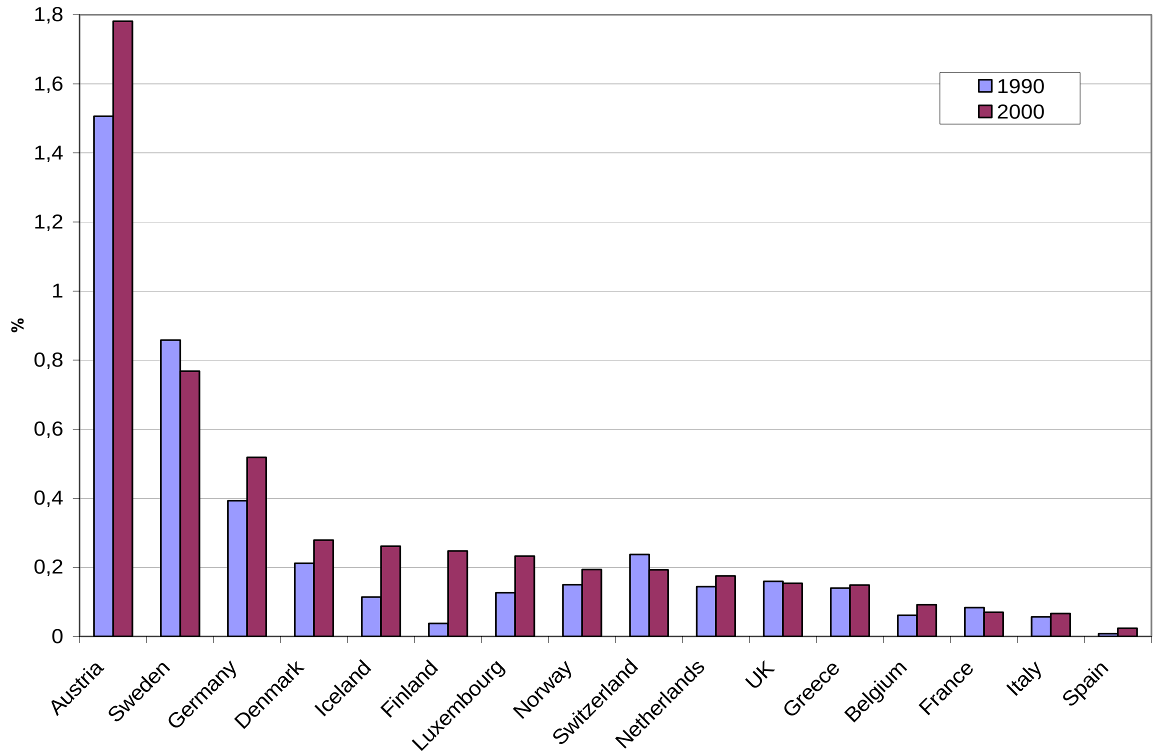
Source: National statistical offices; Own calculations.

Thus, Germany is clearly the most dominant European receiving country[79] of immigrants from the new EU countries in both absolute and relative terms.[80] Further, fairly high migration flows as a proportion of host countries' populations were experienced by Scandinavian countries. In absolute numbers, the Southern European countries like Italy and Spain have also been popular destinations of Eastern Europeans during the latest decade.

[79] However, as can be seen in Pytlikova (2005) there are other popular destinations of Central and Eastern European emigrants outside the EU.
[80] Moreover, the figure does not include Aussiedler, ethnic Germans from the Central and Eastern Europe – if the numbers were included, the share would be much higher.

Although there have been significant flows of immigrants from the new EU members after the breakdown of the Iron Curtain as shown in Figure 1, the stock of immigrants from those countries has remained relatively low at around 0.2% of the population for the vast majority of EEA countries, see Figure 5.2.

Figure 5.2: *Stock of immigrants from the 7 new EU member states residing in the "old" EEA/EU countries as a percentage of the host countries' populations, year 1990 and 2000.*



Note 2: Due to data availability the table shows information on: 1991 instead of 1990 for Austria, Iceland, Italy and Spain; 1991 and 2001 instead of 1990 and 2000, respectively, for Luxembourg; 1999 instead of 2000 for France; 1992 instead of 1990 for the UK. For the UK the figure does not reflect foreigners from Baltic countries as due to confidentiality there is no information on those origins. Ireland and Portugal have been excluded due to missing information on CEE countries of origin.

Source: National statistical offices, Own calculations.

Figure 5.2 shows that the highest percentage of immigrants originating from the new EU countries relative to the host country population is found in Austria reaching 1.8% of the total Austrian population. Here the biggest number is of Czech origin with 54,000 Czechs living in Austria. However, these immigrants did not arrive during the latest decade – the majority of them (90%) came during the period shortly after the Second World War, see Lebhart (2003).

Furthermore, Germany, Sweden and other Scandinavian countries have relatively high percentages of the population coming from the new EU member states. Sweden, which traditionally has been a popular destination country for political refugees from former socialistic countries, experienced a significant outflow of the citizens to their home countries after 1989. This return migration together with a growing overall Swedish population during the nineties contributed to a decline in proportion of the immigrants from the new EU members over the decade 1990-2000, see Figure 5.2. A similar pattern can be observed in the cases of Switzerland and France.

Central and Eastern European countries have an entirely different migration history compared to other typical EU source countries, e.g. Turkey or Southern Europe. Besides the fact that migration flows from those countries were very limited prior to the fall of Iron Curtain, there was also a difference as regards the educational composition of the Central and Eastern European migrants. Those that emigrated from the CEECs as political refugees prior to 1989 had in general very high levels of education. In the period after 1989, migrants have been drawn from other parts of the educational distribution as well, but as the general level of education in those countries is relatively high, the Central and Eastern Europeans still constitute a relatively educated group of migrants.[81] Surveys conducted on migration potential in CEECs confirm this trend into the future as the people that expressed the highest willingness to move are in average relatively well educated and young, see e.g. Fassmann and Hintermann (1997).

Table 5.1 summarizes the stock and flow figures during the 1990s for the 15 EEA/EU countries. The change in the CEE migrant stocks between 1990 and 2000 differs to a great extent from the cumulated flows of CEE immigrants because of out- and return migration.[82] For example, the cumulated CEE migrant inflows to EEA/EU in 1989-2000 were 2.23 million, which is equivalent to 3.10% of the

[81] It is a question, however, how transferable is the education from those countries. Another question is related to the quality of education, especially as regards disciplines within the social sciences that suffered a lot during the forty years of communist regime. For a discussion on this issue see e.g. Ammermueller et al. (2003).

[82] It might also be due to deaths or existing differences in definitions of “migrant”.

source countries' populations from new EU member states, respectively. Nevertheless, the migration stock in 2000 was only 180,000 higher than in 1990.

Table 5.1: Flows and stocks of immigrants from the new EU member countries to 15 EEA/EU countries, absolute number and as a percentage of the source countries' populations, 1990 to 2000.

	Flows from the new EU member states:		
Average annual flow	*Period 1989-2000*		
Absolute	185,856		
% of source population	0.26%		
Total accumulated flow			
Absolute	2,230,268		
% of source population	3.10%		
	Stocks from the new EU member states:		
Stocks	*Year 1990*	*Year 2000*	*change*
Absolute	761,120	941,065	179,945
% of source population	1.06%	1.32%	0.25%

See Note 1 and 2.

Source: National statistical offices, Own calculations.

5.3. THEORETICAL MODEL OF MIGRATION

The theoretical framework for my analysis is built on a modified Hatton (1995) model. The model's foundations lay on traditional "human capital investment" framework (Sjaastad, 1962), but in addition the model incorporates features that help to explain migration dynamics. Specifically, it accounts for the formation of expectations about future utility streams based on past information. Thus, it is possible to distinguish short-run and long-run determinants of emigration.[83] In this paper, I slightly modify the Hatton (1995) model as a theoretical framework for my econometric analysis.[84]

[83] The Hatton (1995) model has been applied by many previous studies, e.g. Fertig (2001), Mitchell and Pain (2003), Zaiceva (2004) and others.

[84] However, the model is still a model of permanent migration. From the previous descriptive part one may think that CEE migration mostly has a temporary character, and therefore a model that would allow for return migration would be more appropriate. Till date no such complex migration

Suppose that a potential migrant *k* maximizes his/her utility and that his/her utility from migration is based on the difference between expected incomes in destination *j* and the potential migrant's country of origin *i*, adjusted for costs of migration, z_{kij}.[85]

The probability of his/her migration is determined by the value of his/her utility from migration. The individual's utility is represented by a concave function[86], in particular, the linear case is considered, i.e.:

$$u_{kij} = \ln E(y_{kj}) - \ln E(y_{ki}) - z_{kij} \tag{5.1}$$

Expected income can be defined as earnings times the probability of finding a job. Thus the model accounts for uncertainty in a Harris and Todaro (1970) fashion:

$$E(y) = we \tag{5.2}$$

where *e* denotes employment rate[87] and *w* real earnings. The expected income in the case of migration is:

$$\ln E(y_{kj}) = \ln(w_{kj}) + \ln(e_{kj}), \tag{5.3}$$

Staying in the home country in general implies a higher probability of finding a job compared with moving abroad where the potential migrant may face employment obstacles as foreigners usually do, e.g. language barriers, difficulties with transferring education and/or qualifications and a lack of information. This can be expressed in expected utility of staying at home in country *i*:

$$\ln E(y_{ki}) = \ln(w_{ki}) + \gamma \ln(e_{ki}), \gamma < 1 \tag{5.4}$$

where γ reflects the fact of a higher possibility of finding a job at the home labor market.

Substitution of (5.3) and (5.4) into (5.1) gives the following form of the difference in expected utility:

$$u_{kij} = \ln(w_{kj}) + \ln(e_{kj}) - \ln(w_{ki}) - \gamma \ln(e_{ki}) - z_{kij}, \gamma < 1 \tag{5.5}$$

---

model does exist due to the fact that it is very hard to derive. The author leaves the challenge for future research. For now I derive the model, which is based on the Hatton (1995) model.

[85] The costs of moving to foreign country may be two-fold: direct out-of-pocket costs of migrating and psychological costs of leaving own country, family and friends.

[86] Assuming risk-averse individual.

[87] The employment rate can be expressed as one minus the unemployment rate, i.e $E(y) = w(1-u)$.

It is reasonable to assume that the potential migrant is a forward-looking individual and therefore he/she takes into account expected future income streams in destination country, *j*, and country of origin, *i*. Therefore the total net present value (NPV) of migrating today is equal to expected future, from time *t+1* on, values of the utility viewed at time t, which can be denoted as u^*_{kijt}. The probability of the individual *k* to migrate from country *i* to country *j* at time *t* is given by:

$$\Pr(M_{kijt} = 1) = \Pr(u^*_{kijt} > 0) \tag{5.6}$$

Assuming an average probability of migration over all individuals *k*, the aggregate emigration rate (m_{ijt}) into a given destination country *j* is thus assumed to be a function of NPV:

$$m_{ijt} = \beta u^*_{ijt} \tag{5.7}$$

where β is an aggregation parameter that measures the impact of utilities on the aggregate migration rate.

The model assumes that a potential migrant forms his/her expectations about the future utility gains on the basis of past experience and information. Further, I assume that the formation of expectations follows a geometric series of values:

$$u^*_{ijt} = \lambda u_{ijt} + \lambda^2 u_{ijt-1} + \lambda^3 u_{ijt-2} + ... + \lambda^S u_{ijt-S+1}, 0 < \lambda < 1 \tag{5.8}$$

where λ is a weight between 0 and 1, i.e. λ^S becomes progressively smaller as *s* increases.[88] The geometric series (8) can be also rewritten as:

$$u^*_{ijt} = \lambda u_{ijt} + \lambda u^*_{ijt-1} \tag{5.9}$$

Substituting (5.9) into (5.7) gives:

$$m_{ijt} = \beta\lambda u_{ijt} + \beta\lambda u^*_{ijt-1} \tag{5.10}$$

Further, assuming that if the equation (5.7) is true for time period *t* then it is true also for time period *t-1*:

$$m_{ijt-1} = \beta u^*_{ijt-1} \tag{5.11}$$

which can be used into (5.10) to eliminate u^*_{ijt-1}:

[88] A lag structure with geometrically declining weights is described as having a Koyck distribution.

$$m_{ijt} = \beta\lambda u_{ijt} + \lambda m_{ijt-1} \tag{5.12}$$

Recalling (5.5), it is possible to rearrange (5.12) into:

$$m_{ijt} = \beta\lambda[\ln(w_{jt}) + \ln(e_{jt}) - \ln(w_{it}) - \gamma\ln(e_{it}) - z_{ijt}] + \lambda m_{ijt-1} \tag{5.13}$$

There are different factors that may affect the costs of migration of a potential migrant, z_{ij}. Generally, the larger the distance between two countries the higher are the direct migration costs associated with transportation. The cultural and linguistic distance is important as well: the more distant the new culture is and the larger the language barrier the higher are the migration costs of an individual to migrate. However, changes and improvements in communication technologies, continued globalization of the economy and declining costs of transportation lead to a decline in direct costs of migration over time. An important role in lowering both types of migration costs, direct and psychological, is played by migration "networks", i.e. networks of family members, friends and people of the same origin that already live in a host country. The "networks" can provide potential migrants with the necessary help and information and thus lower the costs of migration.

Thus, it can be assumed that $\bar{z}_{ijt}$ the migration costs associated with migration from country *i* to country *j* averaged over all individuals *k* are larger with distance, but fall over time and are reduced by the existence of migration networks, which can be expressed as:

$$\bar{z}_{ijt} = \varepsilon_0 - \varepsilon_1 s_{ijt-1} - \varepsilon_2 T + \varepsilon_3 D_{ij}, \varepsilon_1 \geq 0, \varepsilon_2 \geq 0, \varepsilon_3 \geq 0 \tag{5.14}$$

where s_{ijt} denotes a stock of immigrants of *i*-origin residing in country *j* as a proportion of the source country's population[89] at the end *t-1*, *T* is a deterministic linear trend and serves as a proxy for falling migration costs, D_{ij} denotes distance[90] between two countries *ij* and serves as a proxy for growing migration costs with a larger distance. Finally, $\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3$ are coefficients.

The stock of migrants usually serves as a good proxy for migration networks, see Hatton (1995) and Pedersen et al. (2004). The larger the established "networks" of a

[89] The stock of immigrants is expressed as a proportion of the source country's population in order to account for a proportion of networks around potential migrant.
[90] Here, I assume that a distance serves as a kind of spatial trend.

particular ethnic group are, the lower are the migration costs for potential migrants from those ethnic origins. The stock of immigrants is lagged one period as I assume that the networks have to be established at least for some time in order to be effective, i.e. immigrants must have some time to settle down, get orientated in a new country and gain some resources to be able to provide followers with information and other forms of help.

Finally, substituting relations (5.14) into (5.13) gives a model that is used for analyses in this paper:

$$m_{ijt} = a\left[\ln(w_j / w_i)_t + \ln(e_j)_t - \gamma \ln(e_i)_t - \varepsilon_0 + \varepsilon_1 s_{ijt-1} + \varepsilon_2 T - \varepsilon_3 D_{ij}\right] + \lambda m_{ijt-1} \qquad (5.15)$$

where $a = \beta\lambda$. Equation (5.15) is in fact a model in an autoregressive distributed lag model form, which allows distinguishing between the short-run and long-run determinants of international migration flows. In the long run, the steady state has the following form:

$$\overline{m_{ij}} = \frac{a}{1-\lambda}[\ln(w_j / w_i) + \ln(e_j) - \gamma \ln(e_i) - \varepsilon_0 + \varepsilon_1 s_{ij} + \varepsilon_2 T - \varepsilon_3 D_{ij}] \qquad (5.16)$$

However, if we would assume that the potential migrant takes into account not only the expected future values of utility viewed at time *t*, u^*_{kijt}, but also current values of his/her utility, u_{kijt}, then the total NPV of migrating today would be a sum of these two utility values, $NPV = u^*_{kijt} + u_{kijt}$. This introduces the individual's option of waiting[91] with migration into the model. Thus, if the current utility of migration is negative for some individuals, then he/she prefers to postpone the decision to migrate.

Thus the probability of the individual *k* to migrate is given by:

$$\Pr(M_{kijt} = 1) = \Pr(u^*_{kijt} + u_{kijt} > 0 \cap u_{kijt} > 0) \qquad (5.17)$$

Then the aggregate emigration rate in (5.7) is extended to:

$$m_{ijt} = \beta(u^*_{ijt} + \alpha u_{ijt}), \alpha > 1 \qquad (5.18)$$

where α reflects a higher importance given to the current period's utility, given the option to wait with migration.

[91] See Burda (1995) on the option value of waiting in connection to migration.

Holding the prior assumption related to formation of expectations, the emigration rate can be rewritten as:

$$m_{ijt} = \beta(\alpha+\lambda)u_{ijt} - \lambda\beta\alpha u_{ijt-1} + \lambda m_{ijt-1} \tag{5.19}$$

which is possible to rearrange into:

$$\begin{aligned} m_{ijt} &= \beta(\alpha+\lambda)[\ln(w_{jt}) + \ln(e_{jt}) - \ln(w_{it}) - \gamma\ln(e_{it}) - z_{ijt}] \\ &\quad -\lambda\beta\alpha[\ln(w_{jt-1}) + \ln(e_{jt-1}) - \ln(w_{it-1}) - \gamma\ln(e_{it-1}) - z_{ijt-1}] + \lambda m_{ijt-1} \end{aligned} \tag{5.20}$$

Remembering the migration costs equation (5.14), the term z_{ijt} can be substituted by the existing stock of immigrants, by time trend and distance. Further, I assume that the stock diminishes due to death and out-migration and it increases by new waves of immigrants, i.e. there is a change in the migration stock, which makes up a net migration flow. Thus, the stock of immigrants[92] in period *t* is given by the following equation:

$$s_{ijt} = s_{ijt-1} + \Delta s_{ijt} \tag{5.21}$$

Relation (5.21) can be used further to rewrite (5.14) for time period *t-1*[93]:

$$\bar{z}_{ijt-1} = \varepsilon_0 - \varepsilon_1(s_{ijt-1} - \Delta s_{ijt-1}) - \varepsilon_2 T + \varepsilon_3 D_{ij}, \varepsilon_1 \geq 0, \varepsilon_2 \geq 0, \varepsilon_3 \geq 0 \tag{5.22}$$

which is needed to eliminate $\bar{z}_{ijt-1}$ in equation (5.20). Finally, substituting relations (5.14), (5.21) and (5.22) into (5.20) gives:

$$\begin{aligned} m_{ijt} &= \beta(\alpha+\lambda)\ln(w_j/w_i)_t - \lambda\beta\alpha\ln(w_j/w_i)_{t-1} \\ &\quad +\beta(\alpha+\lambda)\ln(e_j)_t - \lambda\beta\alpha\ln(e_j)_{t-1} \\ &\quad -\beta(\alpha+\lambda)\gamma\ln(e_i)_t + \lambda\beta\alpha\gamma\ln(e_i)_{t-1} \\ &\quad -\beta(\alpha+\lambda)\varepsilon_0 + \lambda\beta\alpha\varepsilon_0 \\ &\quad +\beta(\alpha+\lambda)\varepsilon_1 s_{ijt-1} - \lambda\beta\alpha\varepsilon_1 s_{ijt-1} + \lambda\beta\alpha\varepsilon_1\Delta s_{ijt-1} \\ &\quad +\beta(\alpha+\lambda)\varepsilon_2 T - \lambda\beta\alpha\varepsilon_2 T \\ &\quad -\beta(\alpha+\lambda)\varepsilon_3 D_{ij} + \lambda\beta\alpha\varepsilon_3 D_{ij} \\ &\quad +\lambda m_{ijt-1} \end{aligned} \tag{5.23}$$

This expression can be further rearranged into:

[92] The stock of immigrants is expressed similarly as migration flows as a rate, i.e. as a proportion of source country population.
[93] I assume that if the equation (5.14) is true for time period *t* then it is true also for time period *t-1*:

$$\Delta m_{ijt} = \beta(\alpha+\lambda)[\Delta \ln(w_j / w_i)_t + \Delta \ln(e_j)_t - \gamma\Delta \ln(e_i)_t] + \lambda\beta\alpha\varepsilon_1 \Delta s_{ijt-1}$$
$$+\beta(\alpha+\lambda-\lambda\alpha)[\ln(w_j / w_i)_{t-1} + \ln(e_j)_{t-1} - \gamma \ln(e_i)_{t-1} + \varepsilon_1 s_{ijt-1} + \varepsilon_2 T - \varepsilon_3 D_{ij} - \varepsilon_0] \quad (5.24)$$
$$-(1-\lambda)m_{ijt-1}$$

Equation (5.24) presents a dynamic model in an error correction form. Here, the long-run steady state is given by:

$$\overline{m_{ij}} = a_1[\ln(w_j / w_i) + \ln(e_j) - \gamma \ln(e_i) + \varepsilon_1 s_{ij} + \varepsilon_2 T - \varepsilon_3 D_{ij} - \varepsilon_0] \quad (5.25)$$

where $a_1 = \dfrac{\beta(\alpha+\lambda-\lambda\alpha)}{(1-\lambda)}$ and where the parameters do not depend on time.

Then the models in the forms (5.15) and (5.24) can be used further as a base for econometric analyses.

5.4. ECONOMETRIC ANALYSIS

5.4.1. Data

For the analysis I use information on migration flows and stocks in 13 EEA/EU destination countries from the 7 new EU member source countries for the years 1989-2000, see Appendix 5.IIA for a list of countries considered in the analysis. I have collected the migration information by contacting statistical offices of particular OECD destination countries.[94]

Besides the actual migration flow and stock information, the dataset contains a number of other time-series variables, which may help to explain the determinants of migration across countries. For purposes of the current paper, only information on GDP per capita, unemployment rates, population and distance has been used. These variables were collected from different sources, e.g. OECD, the World Bank and others; see Appendix 5.IIB for definitions and sources of the variables. For a more comprehensive description of the dataset, see Pedersen et al. (2004, 2005).

[94] Originally, the dataset has been constructed for 27 OECD destination and 129 source countries, see Pedersen et al. (2004) for a detailed description of the dataset. In this paper, I restricted it to a sample of 13 destination and 7 CEE source countries. Destination countries as Austria and France have been excluded from the analyses due to the fact that there are just two observations of the stock of immigrants from censuses. This makes it impossible to create sensible time-series for these countries.

Although the dataset presents a large progress over the datasets used in earlier East-West migration research, there are still some problems. First, the dataset is unbalanced, i.e. observations are missing in the panel. For the majority of destination countries, I have information on migration flows and stocks of immigrants for most of the years, but often with different numbers of observation for each destination country. Another problem is that different countries use different definitions of an "immigrant"[95] and different sources for their migration statistics. For a description of some of the difficulties related to collecting international migration data, see Pedersen et al. (2004, 2005). For summary statistics, see Appendix 5.IIC.

5.4.2. Choice of preferred econometric specification - model and estimator

5.4.2.1. Static model

In this section, I present econometric analyses of migration potential from new EU member countries. I base my econometric analysis on the human capital model presented in the previous section. The model assumes that emigration is driven by relative mean earnings and employment rates between origins and destinations, and the costs of migration as captured by existing networks of migrants, distance and a time trend. I start with the econometric specification, which has the following form:

$$\begin{aligned} \ln m_{ijt} &= \gamma_1 + \gamma_2 \ln(GDP_j / GDP_i)_{t-1} + \gamma_3 \ln e_{jt-1} + \gamma_4 \ln e_{it-1} + \gamma_5 \ln s_{ijt-1} \\ &+ \gamma_6 dist_{ij} + \gamma_7 T + \mu_{ij} + \varepsilon_{ijt} \end{aligned} \tag{5.26}$$

where m_{ijt} denotes gross flows of migrants from country *i* to country *j* divided by the population of the country of origin *i* at time *t*[96], where *i=1,…,7; j=1,…,13* and *t=1,...,11.* The difference in earnings is approximated by relative differences in economic development measured by GDP per capital in PPP and enters the model as a ratio, GDP_j / GDP_i. The employment opportunities in the sending and receiving countries measured by employment rate (1-unemployment rate) are denoted as e_j and e_i, respectively. The network links between sending and receiving countries that

[95] Some use foreign-born (by countries of birth), some use foreigners (by citizenship/nationality).
[96] I estimate the model with net migration rates on the left-hand side as well, but I return to that later on in the paper.

help to lower the costs of migrating are captured by the normalized stock of immigrants, s_{ij}, i.e. the stock of immigrants from source country *i* living in destination *j*, divided by population in source country *i*. Variable $dist_{ij}$ denotes a distance in kilometers between two countries, which serves as a proxy for the direct costs of migration. The declining travel costs of migration over time are captured by the trend variable T. Finally, μ_{ij} denotes unobservable country-specific effects and ε_{ijt} denotes remaining error. I use all variables (except the trend) in logs to show the impact elasticities. In this econometric model, I omit the lagged dependent variable that enters the theoretical model in the previous section. I will come back to the specification with the dynamic term in the next section.

From the economic theory point of view, the relative differences in economic development and employment should be lagged in order to account for the information, on which the potential immigrants base their decision to move. Further, there might be a reverse causality with respect to the effect of migration flows on earnings and employment.[97] One way to avoid the problems of endogeneity in the model is to instrument earnings and employment variables with their lags. As regards the migrants' network, the variable is endogenous, too, as in fact the stock is a function of previous stock plus migration flows minus out-migration. Therefore, all the explanatory variables enter the model as lagged. According to the theory, I expect estimated coefficients to have the following signs:

$\gamma_2 > 0, \gamma_3 > 0, \gamma_4 < 0, \gamma_5 > 0, \gamma_6 < 0$, and $\gamma_7 > 0$.

Besides the explanatory variables covered in the model above, there are, naturally, other variables that can help to explain migration behavior and that might be included in the econometric analysis.[98] But this paper focuses on the present simplified specification for several reasons. First, the model is used primarily for the purpose of predictions of the CEE migration potential. As it is relatively difficult to make reliable predictions of other explanatory variables themselves, the model should stay as uncomplicated as possible. Second, the model above is typically used

[97] There is another huge stream of literature that focuses on the effect of immigration on the labor market, see e.g. Chiswick (1996), Filer (1992), Hunt (1992) and Chiswick and Hatton (2002).

[98] For instance variables capturing language, cultural barriers, education, trade and other, see e.g. Karemera et al. (2000), Pedersen et al. (2004, 2005), Pytlikova (2005), Belot and Ederveen (2005) and Mayda (2005) for discussions on determinants of migration.

in previous studies assessing East-West migration potential. Thus, it is possible to compare my results with the ones from the previous studies. For these reasons, I stick to the explanatory variables suggested by the theoretical model.

As a starting point of my analysis, I estimate the model (26) above by using pooled OLS,[99] see Table 5.2. As I deal with international migration, one would suspect an existence of unobserved country-specific heterogeneity. In this case, the pooled OLS estimator is biased and inconsistent, see Baltagi (2005), and therefore the fixed and random effects panel data estimators are estimated in the next step. The rich structure of my dataset gives me an opportunity to evaluate whether to consider unobservable effects of destinations, origins or pairs of those countries. One could interpret e.g. destination country fixed effects, FE(j), as destination's climate, weather, openness towards foreigners, culture, existence of widely spoken language, or immigration policy, which do not change over time and which is common for all countries of origin. On the other hand, unobservable pairs of countries fixed effects, FE(ij), capture traditions, historical and cultural ties between two particular countries of destination and origin, as well as the immigration policy between the two countries. According to the adjusted R-squared in Table 5.2, the pairs of countries unobserved effects panel data estimator would be preferred.

At the bottom of Table 5.2, the F-test and Breush and Pagan LM test show that both fixed and random effects are significant, and thus the pooled OLS that omits the country-specific effects indeed suffers from an omission variable bias. In order to choose a suitable panel data estimator, I evaluate the relationship between the country–specific effects and the explanatory variables by using a Hausman specification test. According to the Hausman test, the fixed effects estimator is preferred over the random-effects estimator in all specifications except in the destination countries' unobserved effects estimators. Overall, taking into account the econometric issues discussed above, the most preferred estimator for the static model is the pairs of countries fixed effects estimator, FE(ij), shown in column 4 in Table 5.2.

[99] I estimate the model by OLS with robust - Huber/White/sandwich – estimate of variance and I also use a cluster option for pair of countries in order to adjust standard errors for intra-group correlation.

Table 5.2: Estimation of migration flows from 7 source countries (i) to 13 EEA/EU destination countries (j), 1990-2000.

Dependent variable	Gross flows per source country population $\ln m_{ijt}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	FE(j)	RE(j)	FE(ij)	RE(ij)	FE(i)	RE(i)
Independent variables:							
Migration stock, $\ln(s_{ij})_{t-1}$	0.572 [0.038]***	0.374 [0.036]***	0.386 [0.035]***	0.278 [0.048]***	0.364 [0.042]***	0.643 [0.032]***	0.572 [0.032]***
lnGDP ratio *t-1*	1.444 [0.209]***	2.287 [0.179]***	2.242 [0.179]***	1.012 [0.354]***	0.747 [0.284]***	0.194 [0.279]	1.444 [0.214]***
Employment rate *j*, $\ln(e_j)_{t-1}$	-1.561 [1.296]	8.671 [1.668]***	7.257 [1.609]***	9.032 [1.121]***	7.397 [1.085]***	0.373 [1.413]	-1.561 [1.423]
Employment rate *i*, $\ln(e_i)_{t-1}$	1.873 [1.394]	4.238 [0.951]***	4.196 [0.955]***	0.058 [0.820]	0.277 [0.804]	2.065 [1.545]	1.873 [1.391]
lnDistance	-0.830 [0.120]***	-0.982 [0.102]***	-0.961 [0.102]***	-	-1.315 [0.254]***	-0.695 [0.120]***	-0.830 [0.120]***
Trend	0.046 [0.021]**	0.091 [0.014]***	0.090 [0.014]***	0.060 [0.011]***	0.052 [0.011]***	0.028 [0.021]	0.046 [0.021]**
Fixed/Random Effects of Destination	No	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	591	591	591	591	591	591	591
No. of destinations / origins/ pairs of countries		13	13	78	78	7	7
Constant Term Included	-0.203 [7.362]	-57.672 [8.214]***	-50.998 [7.959]***	-46.061 [5.794]***	-29.924 [5.726]***	-9.265 [8.046]	-0.203 [7.343]
Adj. R-square	0.64	0.86	0.55	0.94	0.60	0.69	0.64
RMSE	1.207	0.749	0.749	0.450	0.454	1.108	1.201
F- test		F(12,572) =74.99, p=0.000		F(77, 508) =40.36, p=0.000		F(6,578) =16.72, p=0.000	
Breush-Pagan test		chi2(1) =4696.92, p=0.000		chi2(1) =1360.89, p=0.000		chi2(1) =177.73, p=0.000	
Hausman test		chi2(6) =10.37, p=0.110		chi2(5) =54.67, p=0.000		chi2(6) = 125.15, p=0.000	

Notes: 10, 5 and 1% levels of confidence are indicated by *, ** and ***, respectively. Standard errors are in parentheses.

When looking at results across the different estimators, they seem to be quite robust over all specifications. Concentrating on the results from my preferred FE(ij) estimator, the coefficients have the expected signs except the employment rates in the source countries. Income differences and employment prospects in destinations appear to be key driving forces behind the CEE emigration. The source employment rate variable has an unexpected positive sign, but it is insignificant. Such an effect is often found in other studies on the determinants of international migration, e.g. Faini and Venturini (1994), Hatton and Williamson (2002) and Pedersen et al. (2004). It might be explained by poverty constraints of potential migrants, i.e. higher employment gives necessary resources to emigrate. In the case of Central and Eastern European countries besides the poverty constraints, it might be due to the exceptional development in unemployment during the first years of transition.

5.4.2.2. Dynamic model

In the static model specification (5.26), I have omitted the dynamic terms that are in fact extensively discussed in the theoretical model. Specifically, the dynamics are introduced into the model by adding a lagged dependent variable.[100] This allows for modeling of persistence in migration caused by the outcomes of previous periods in addition to the persistence caused by the unobserved country-specific effects. Thus, short-run and long-run effects are possible to distinguish. The dynamic econometric model can be presented in two different forms: in the form of an autoregressive distributed lag model (ARDL) or in the form of an error-correction model (ECM), which are equivalent to the equation (5.15) and (5.24) from the theoretical model.

First, the ARDL[101] model of migration can be written as:

$$\begin{aligned} \ln m_{ijt} &= \beta_0 + \beta_1 \ln m_{ijt-1} + \beta_2 \ln(GDP_j / GDP_i)_{t-1} + \beta_3 \ln e_{jt-1} + \beta_4 \ln e_{it-1} + \beta_5 \ln s_{ijt-1} \\ &+ \beta_6 dist_{ij} + \beta_7 T + \mu_{ij} + \varepsilon_{ijt} \end{aligned} \tag{5.27}$$

The short-run elasticities are represented by estimates of $\beta_2, \beta_3, \beta_4, \beta_5, \beta_6$ and β_7. The long-run elasticities for each variable are calculated as:

$$\frac{\beta_2}{1-\beta_1}, \frac{\beta_3}{1-\beta_1}, \frac{\beta_4}{1-\beta_1}, \frac{\beta_5}{1-\beta_1}, \frac{\beta_6}{1-\beta_1} \text{ and } \frac{\beta_7}{1-\beta_1} \tag{5.28}$$

Next, I consider the dynamic migration model in an error-correction form, which can be written as follows:

$$\begin{aligned} \Delta \ln m_{ijt} &= \delta_0 + \delta_1 \Delta \ln(GDP_j / GDP_i)_t + \delta_2 \Delta \ln e_{jt} + \delta_3 \Delta \ln e_{it} + \delta_4 \Delta \ln s_{it-1} \\ &+ \phi EC_{ij} + \mu_{ij} + \varepsilon_{ijt} \end{aligned} \tag{5.29}$$

[100] The dynamics enter the equation also through the migration stock variable, which contains previous history of migration flows. To be precise, the stock of immigrants consists of the previous stock plus the migration in-flows minus the migration out-flows.
[101] In fact, it is an ARDL (1,1) model.

The parameters to the first difference denoted by Δ measure the short-run effects. The parameter to the error-correction component, ϕ, is interpreted as a speed of adjustment to the long-run equilibrium relationship.[102] A small value of the error-correction component indicates that any shock to aggregate migration activity typically has a long lasting impact.

The error-correction model is estimated in a two-step procedure. First, the long-run parameters are obtained from the long-run migration relationship:[103]

$$\ln \bar{m}_{ij} = a_0 + a_1 \ln(GDP_j / GDP_i) + a_2 \ln(e_j) + a_3 \ln(e_i) + a_4 \ln(s_{ij}) + a_5 dist_{ij} \tag{5.30}$$

The error-correction component, *EC*, can then be calculated from:

$$EC = \ln m_{ijt-1} - \hat{a}_1 \ln(GDP_j / GDP_i)_{t-1} - \hat{a}_2 \ln(e_j)_{t-1} - \hat{a}_3 \ln(e_i)_{t-1} - \hat{a}_4 s_{ijt-1} - \hat{a}_5 dist_{ij} - \hat{a}_0 \tag{5.31}$$

And finally the full model as given in (5.29) is estimated.

The dynamic models can be estimated by different panel data techniques. Once analyzing the migration behavior in the static framework, I found that pairs of country-specific effects play an important role in explaining migration flows. Therefore, these unobservable effects should be included in the dynamic modeling, too.

One potential problem with the fixed or random panel data techniques in an ARDL framework is that both estimators are biased and inconsistent once the lagged dependent variable enters the model, especially in short panels like the one used in this study, see e.g. Baltagi (2005). It is possible to tackle the problem by using the Arellano-Bond (1991) difference (diffGMM) or the Arellano-Bover (1995) system

---

[102] From the microeconomic behavior point of view, the error correction term accounts for the existence of uncertainty and imperfect information connected to migration and the option value of waiting as argued in the previous theoretical model section.

[103] Here, the condition of co-integrating relationship must be satisfied. In order to test whether the co-integrating relationship exists, I first tested for existence of unit roots in the variables by using the Fisher's panel unit root test. The results are rather mixed and reveal that there are specific co-integration relations for each pair of countries. However, given the short panel of the data, one should be careful with any strong interpretation from the unit root tests run on the data. For a detailed description of the testing techniques and the results of panel unit roots tests, see Appendix II D. Further, I apply the augmented Dickey-Fuller (pooled) test in order to detect the stationarity in the error terms of the long-run migration equation. The test supports the hypothesis that the variables above form a co-integrating relationship; see bottom of Table 5.3. But as the test is not a panel data test, the results should be treated with caution.

GMM estimator (sysGMM). The first one uses instrumental variables in a form of lagged levels of predetermined and endogenous variables in first-differenced equations. The system GMM additionally permits the original equations in levels to be added into the model and predetermined and endogenous variables in levels to be instrumented by their suitably lagged first differences.[104] In this paper, both difference and system GMM estimators are performed in a two-step set up. I acknowledge possible heteroscedasticity by presenting robust standard errors.[105]

Results of the ARDL and ECM dynamic models are presented in Table 5.3. The first four columns show the ARDL model estimated by simple pooled OLS[106], countries pair's fixed effect FE(ij), system and difference GMM two-step estimators, (sysGMM, diffGMM).[107]

[104] For endogenous variables, I used lagged levels dated t-2 as instruments in first difference equations and lagged first-differences dated t-1 as instruments in level equations. For predetermined variables, their lagged levels dated t-1 and their first differences are used as instruments in first difference and level equations, respectively. I treat lagged migration flows and migration stock as endogenous variables, the GDP ratio and employment rates in destination and origin as predetermined, and trend and distance as exogenous. A precise list of the instruments is given in the Note of Table 5.3.

[105] As the panel is short, the Windmeijer's finite sample correction has been applied. Both estimators have been conducted using the "xtabond2" command developed independently by Roodman (2005) for STATA9.

[106] With a robust estimate of variance.

[107] I test the validity of instruments by the Hansen test of over-identifying restrictions, which tests whether the instruments as a group appear exogenous. As can be seen from Table 5.3, the Hansen test of over-identifying restrictions cannot be satisfied for both GMM estimators in a two-step set up. However, the Hansen test turns out to be satisfied in a one-step set up, which shows a sensitivity of the test to different set ups. The results from the one-step GMM estimations are available from the author upon request. The sufficiency of GMM is further tested by autocorrelation tests derived by Arellano and Bond (1991). For both GMM estimators the first-order serial correlation test rejects the null hypotheses of no first-order serial correlation (Arellano-Bond AR (1) test: z = -2.14 and -2.62), whether the second-order serial correlation test is unable to reject the null hypotheses of no second-order serial correlation (Arellano-Bond AR (2) tests: z = 0.63 and 1.04).

Table 5.3: Estimation of gross migration flows in the dynamic ARDL and ECM framework from 7 new EU member countries to 13 EEA/EU countries, 1990-2000.

	ARDL Model				Error Correction Model	
Dependent variable	Gross migration rate $\ln m_{ijt}$			Change in gross migration rate $\Delta \ln m_{ijt}$		
	(1)	(2)	(3)	(4)	(5)	(6)
Independent variables:	OLS	FE(ij)	sysGMM (ij) two-step	diffGMM (ij) two-step	OLS	FE(ij)
Short-run	*In levels*				*In changes*	
Migration stock, $\ln(s_{ij})_{t-1}$	0.026 [0.022]	0.112 [0.053]**	0.056 [0.067]	0.542 [0.105]***	0.077 [0.065]	-0.108 [0.067]
lnGDP ratio *t-1*	0.280 [0.092]***	0.342 [0.365]	0.632 [0.221]***	0.100 [0.424]	1.334 [0.551]**	1.790 [0.533]***
Employment rate *j*, $\ln(e_j)_{t-1}$	-0.358 [0.683]	6.949 [1.120]***	-0.130 [0.933]	8.096 [1.633]***	6.390 [1.752]***	1.752 [1.699]
Employment rate *i*, $\ln(e_i)_{t-1}$	0.778 [0.690]	-0.546 [0.790]	1.847 [0.719]**	0.189 [0.992]	1.496 [0.869]*	0.719 [0.823]
Distance	-0.050 [0.054]	-	-0.103 [0.100]	-	0.058 [0.039]	-
Trend	0.028 [0.010]***	0.042 [0.011]***	0.025 [0.009]***	-0.009 [0.016]	-	-
Long-run ♠						
Migration stock, $\ln(s_{ij})$	0.333	0.165	0.397	0.654	0.620 [0.030]***	
GDP ratio	3.590	0.503	4.482	0.121	1.192 [0.196]***	
Employment rate *j*, $\ln(e_j)$	-4.590	10.219	-0.922	9.766	-1.666 [1.314]	
Employment rate *i*, $\ln(e_i)$	9.974	-0.803	13.099	0.228	0.545 [1.232]	
Distance	-0.641	-	-0.730	-	-0.764 [0.110]***	
Trend	0.359	0.062	0.177	-0.011	0.031 [0.018]*	
Dynamic components						
Error-correction term	-	-	-	-	-0.078 [0.020]***	-0.574 [0.047]***
Migration flows, $\ln(m_{ij})_{t-1}$	0.922 [0.022]***	0.320 [0.046]***	0.859 [0.052]***	0.171 [0.091]*	-	-
Fixed/Random Effects - Country Pair	No	Yes	Yes	Yes	No	Yes
No. of observations	561	561	597	519	547	547
Constant Term Included	-2.202 [4.127]	-32.118 [5.822]***	-8.168 [5.928]	no	-0.320 [0.275]	0.062 [0.024]***
Adj. R-square	0.93	0.95			0.07	0.27
RMSE	0.532	0.406	0.557	1.124	0.522	0.429
Hansen test			chi2(316) = 70.3 p=1.00	chi2(264) = 70.51 p=1.00		
Dickey-Fuller test					chi2(150) = 381.99, p=0.00	

Notes: 10, 5 and 1% levels of confidence are indicated by *, ** and ***, respectively. Standard errors are in parentheses. Following instruments were used for diffGMM (column 4): ln migraton rate t-2, ln stock t-2, ln GDP ratio t-1, ln employment j t-1, ln employment i t-1and all further lags. Additional instruments used for levels equations in sysGMM (column 3): Δ ln migration flows t-1, Δ ln stock t-1, Δ lnGDP ratio t, Δ employment j t, Δ employment i t and all further lags. ♠ In ARDL models, the long-run coefficients are calculated from short-run elasticities and lagged dependent variable as described in (5.28).

The last two columns in Table 5.3 contain results from the error-correction model estimated by using OLS and the countries pair's fixed effects FE(ij) panel data estimator. Similarly as in the model without the lagged dependent variable, the unobserved country pairs fixed effect estimator in the ARDL framework gives the best performance as measured by the RMSE. The FE(ij) in ECM framework and sysGMM perform relatively well, too.

As regards results, the coefficients mostly attach the expected signs, but their size varies across the different econometric specifications. The coefficients to the dynamic component have the correct signs and are significantly different from zero, which confirms that statistically valid long-term equilibrium exists, and thus migration rates indeed adjust towards some long-run level. But, most importantly, the size of the coefficient varies significantly across the different specifications. To be precise, the coefficient to the lagged dependent variable and to the error-correction term in the models estimated by OLS or sysGMM, see columns 1, 3 and 5 in Table 5.3, have relatively lower magnitude compared to the models estimated by FE(ij) and diffGMM estimators, see columns 2, 4 and 6 in Table 5.3. This suggests that the adjustment in migration rate towards its equilibrium would be more sluggish according to the former ones and relatively faster according to the latter mentioned estimators.

The relative lack of robustness across different estimators in the dynamic models with lagged dependent variable on the right-hand side might be due to the nature of the panel data set used for the analysis. In particular, its unbalanced structure and/or its short-time panel dimension may cause the instability of the results in the dynamic models.

5.4.2.3. Preferred econometric specification to be used for predictions

The different estimators both in the static and dynamic frameworks, Tables 5.2 and 5.3, show that it is important to control for pairs of countries unobserved heterogeneity in the intercepts. The fixed country pair's effects estimator in its dynamic form, see column 2 in Table 5.3, beats the other estimators on the basis of explanatory power measured by the adjusted R-squared and in the terms of the RMSE. Therefore, this estimator is preferred in my analysis and is used for the predictions of a future migration potential.

Further, in addition to the gross migration rate as the dependent variable, I estimate similarly - as done in other studies - the model (25) with the net migration rate nm_{ijt} on the left-hand side:

$$\begin{aligned} nm_{ijt} &= \gamma_1 + \gamma_2 nm_{ijt-1} + \gamma_3 \ln(GDP_j / GDP_i)_{t-1} + \gamma_4 \ln e_{jt-1} + \gamma_5 \ln e_{it-1} \\ &+ \gamma_6 \ln s_{ijt-1} + \gamma_7 dist_{ij} + \gamma_8 T + \mu_{ij} + \varepsilon_{ijt} \end{aligned} \tag{5.32}$$

The net migration flows numbers are obtained from the change in stocks per source population: $nm_{ijt} = \Delta s_{ijt} = s_{ijt} - s_{ijt-1}$ and enter the equation in a non-log form.[108] The coefficients from the model (5.32) are estimated similarly as in the gross migration flows equation by using the fixed effects FE_{ij} panel data estimator.[109]

The results of the models to be used for the predictions of the gross and net migration flows are presented in Table 5.4.

Overall, the two regressions show different effects of the explanatory variables on the gross and net migration flows. For instance, the coefficient to the migration stock attaches an unexpected negative sign in the net flows regression. The similar result is found in e.g. Fertig (2001) who argues that the stocks variable captures not only the "network" effect, but also "decreasing returns to migration". The larger stocks of immigrants imply a harder competition on the labor market, which leads to larger migration outflows. Further, the coefficient to the lagged GDP ration has an unexpected negative sign. This has a direct implication for the carried predictions as the predicted net migration flows will increase with the higher GDP convergence, but will be decreasing slowly with the growing stocks of immigrants. The lagged dependent migration rates yield statistically significant positive coefficients in both regressions. This means that a long-run equilibrium exists for both gross and net migration rates. Finally, the coefficients to the employment rates in the destinations and origins have expected signs in both gross and net migration rate regressions.

[108] The reason why the model is in the semi-log form and not as the model with gross migration rates in the log-log form is that there are some negative values of the net migration rate for some pairs of countries, e.g. Czechs and Hungarians in Germany in 1995-1998 and 1995-1997, respectively, Hungarians in Switzerland in 1992-1996, Estonians in Sweden in 1996-1999, Czechs, Hungarians and Poles in the UK for some years. Therefore, a model in semi-log form is more suitable.

[109] Results of the model estimated by other econometric specifications: OLS, REij, diffGMM and sysGMM are available from the author upon request. The FE_{ij} is preferred over the RE_{ij} in the Hausman specification test.

Table 5.4: Fixed effects FE_{ij} estimation of gross and net migration flows from 7 new EU members to 13 EEA/EU countries, 1990-2000.

	New EU member countries	
Dependent variable:	Gross flows per source country population, $\ln(m_{ij})_t$	Net flows per source country population, nm_{ijt}
Independent variables:		
Gross migration flows, $\ln(m_{ij})_{t-1}$	0.320 [0.046]***	
Net migration flows, nm_{ijt-1}		0.378 [0.038]***
Migration stock, $\ln(s_{ij})_{t-1}$	0.112 [0.053]**	-0.026 [0.013]*
lnGDP ratio *t-1*	0.342 [0.365]	-0.192 [0.094]**
Employment rate *j*, $\ln(e_j)_{t-1}$	6.949 [1.120]***	0.280 [0.307]
Employment rate *i*, $\ln(e_i)_{t-1}$	-0.546 [0.790]	-0.121 [0.218]
Trend	0.042 [0.011]***	-0.002 [0.003]
Fixed/Random Effects of Country Pair	Yes	Yes
No. of observations	561	607
No. of pairs of countries	78	76
Constant Term Included	-32.118 [5.822]***	-0.557 [1.585]
Adj. R-square	0.95	0.39

Notes: 10, 5 and 1% levels of confidence are indicated by *, ** and ***, respectively. Standard errors are in parentheses.

5.5. PREDICTIONS OF MIGRATION POTENTIAL

In the following section, I carry out a prediction exercise based on the models that have been presented in Table 5.4. Specifically, I predict the gross and net migration flows from the 7 new EU countries over the period 2004 to 2015. In connection with that, I make several assumptions concerning the future developments of the explanatory variables.

First, as regards the development of GDP per capita and the convergence between the source and destination countries, I assume, as Alvarez-Plata et al. (2003), three different scenarios:

a) *High convergence scenario* with a 3 percentage points convergence rate. The source countries are assumed to grow 5% annually where the EEA/EU destination countries are assumed to grow 2% annually;

b) *Medium convergence scenario* with a 2 percentage points convergence rate. The source countries are assumed to grow 4% annually where the EEA/EU destination countries are assumed to grow 2% annually;

c) *Low convergence scenario* with a 1 percentage points convergence rate. The source countries are assumed to grow 3% annually where the EEA/EU destination countries are assumed to grow 2% annually.

Further, I assume the population size and the employment rates in both destinations and origins to be stable over time.[110] Thus, I take their values from year 2000 to be the same over the projection period 2004-2015.

As to the network effects, I model the development of migration stock endogenously, i.e. first carrying out the prediction of the net migration flows and then adding the predicted net migration flows to the migration stock from the previous year: $\hat{s}_{ijt} = s_{ijt-1} + n\hat{m}_{ijt}$. Next, the acquired stock value is used to predict the net migration flows in time t+1, $n\hat{m}_{ijt+1}$. I continue with such recursive estimations up to the year 2015. Finally, I also use the obtained migration stock in the model with the gross migration rates as a dependent variable. In this way, I do not have to base my predictions of the migration flows on time-invariant migration stock values as it has been assumed in some previous studies on migration potential, e.g. Hille and Straubhaar (2001), Zaiceva (2004).

Concerning the unobserved pair of countries specific effect, I have first obtained an estimate of the effect specific for each pair of destination and CEE source countries for both the net and gross migration regression models, separately. These (time constant) fixed effects are included in the predictions. The results from this prediction exercise for gross flows and net migration flows from the 7 new EU countries are shown in Tables 5.5 and 5.6, respectively.

Before the description of results, there are a few methodological/data remarks. First, for Spain and Italy, the net migration rates show the stock of immigrants originating from Poland only. For the UK, the figures on stock variables do not cover Baltic

[110] From the previous sections one may observe that there have been drastic declines in the fertility rates in the CEECs. Nevertheless, it is relatively hard to make precise predictions on the population growth. Therefore I assume it to be stable over time.

countries: Estonia, Latvia and Lithuania.[111] Lastly, the UK was dropped from the gross migration flow analysis because the flow data are based on application for settlement statistics only. Consequently, the flow numbers constitute just a fraction of the true migration flow from the new EU members to the UK. I present only predictions of the net flows and stocks from those countries to the UK, as the migration stock statistics is the only reliable source of migration records for the UK. All the details related to the data in connection with the predictions are described in notes under Tables 5.5, 5.6 and 5.7.

Returning to the prediction results, I begin with the estimates of the predicted *gross* migration flows in Table 5.5. Assuming the medium convergence scenario, the average gross flows amount to an average annual immigration of around 322,000 persons from the new EU member countries, which is equivalent to accumulated gross flows of 3.9 million people from those countries or 5.6% of the source countries' populations over the period 2004-2015. This result belongs to the high end of the forecasts from the previous literature.[112] The highest flows from the new EU members are obviously expected to head towards Germany with the average gross migration flow being around 275,000 CEE migrants annually.

[111] The study by Portes and French (2005) shows that the (gross) migration flows from Baltic countries to the UK were especially significant after the opening of the UK labor market for new EU members in 2004.

[112] The results are comparable to two recent studies by Bruder (2003) and Zaiceva (2004). The first study predicts the total accumulated gross flows from CEECs over the period 2004-2015 at around 4.2% of the source countries' populations. The latter study predicts the accumulated gross migration flows between 3 and 5%, see Appendix Table 5.IA for an overview of results of the existing studies.

Table 5.5: Predicted gross migration flows from the new EU members into the 13 EEA/EU countries, 2004-2015.

Gross Flows	**1989-2000**	**Year 2000**	**High convergence scenario 2004-2015**			**Medium convergence scenario 2004-2015**			**Low convergence scenario 2004-2015**		
Countries:	Yearly average		Yearly average	cumulative	% of source	Yearly average	cumulative	% of source	Yearly average	cumulative	% of source
Belgium	1510	3659	4213	50559	0,071	4238	50859	0,071	4226	50712	0,071
Denmark	1472	2515	5189	62266	0,087	5257	63085	0,088	5290	63483	0,089
Finland	1536	1024	4472	53660	0,075	4623	55477	0,078	4777	57318	0,080
Germany	154625	117102	263887	3166649	4,424	274494	3293933	4,602	285688	3428260	4,790
Greece	1092	577	1607	19285	0,027	1610	19319	0,027	1597	19163	0,027
Italy*	3673	7100	10062	120747	0,312	10395	124735	0,323	10734	128810	0,333
Netherlands	1926	2574	6293	75511	0,105	6452	77419	0,108	6599	79186	0,111
Spain**	953	4331	3949	47391	0,087	4069	48830	0,090	4186	50227	0,092
Sweden	1819	1493	3810	45717	0,064	3925	47103	0,066	4039	48467	0,068
EU-9	**168606**	**140244**	**303482**	**3641786**	**5,253**	**315063**	**3780760**	**5,453**	**327135**	**3925626**	**5,660**
Iceland	242	497	700	8403	0,013	680	8160	0,012	655	7857	0,012
Norway	734	859	1811	21737	0,030	1805	21659	0,030	1779	21347	0,030
Switzerland	1547	2129	4235	50821	0,071	4318	51811	0,072	4383	52599	0,073
EEA/EU-12	**171129**	**143729**	**310228**	**3722746**	**5,366**	**321866**	**3862390**	**5,567**	**333952**	**4007429**	**5,775**

Notes: The UK is excluded due to the fact that gross migration flows are based on the settlement migration definition.[113] *For Italy, only numbers of Poles are shown. ** For Spain, only migration from Poland and the Czech Republic is shown. Figure IIIA in the Appendix III shows prediction intervals for the mean gross migration rate across the three different convergence scenarios. Prediction errors for each pair of countries are available from the author upon request.

There might be observed rather lower migration tendencies in the development of the predicted *net* flows from the new EU member countries. Table 5.6 shows that the predicted *net* flows amount to an annual average of 33,000 CEE migrants only. This is equivalent to an accumulated net increase of the CEE residents by 391,000 or by 0.6% of the source countries' populations in the 13 EEA/EU countries over the period 2004-2015.

[113] Consequently the predicted gross migration rates are heavily underestimated. The results from the predictions show an average yearly gross migration flow of 2 940, which makes an accumulated flow over the period 2004-2015 equal to 35284 CEE migrants.

Table 5.6: Predicted net migration flows from the new EU members into the 13 EEA/EU countries, 2004-2015.

Net Flows	**Year 2000**	**High convergence scenario 2004-2015**			**Medium convergence scenario 2004-2015**			**Low convergence scenario 2004-2015**		
Countries:		Yearly average	Cum. increase	% of source	Yearly average	Cum. increase	% of source	Yearly average	Cum. increase	% of source
Belgium	583	2188	26255	0,037	1171	14053	0,020	308	3692	0,005
Denmark	438	2878	34531	0,048	1716	20587	0,029	665	7975	0,011
Finland*	390	1518	18211	0,030	1075	12902	0,021	819	9822	0,016
Germany	16827	13604	163251	0,228	11972	143665	0,201	10325	123899	0,173
Greece	6	1976	23707	0,033	959	11512	0,016	90	1085	0,002
Italy**	4459	3915	46978	0,122	3094	37125	0,096	2271	27250	0,071
Netherlands	1588	4113	49362	0,069	2758	33092	0,046	1445	17340	0,024
Spain**	1626	1936	23231	0,060	1228	14738	0,038	549	6586	0,017
Sweden	362	3993	47917	0,067	2545	30542	0,043	1117	13405	0,019
UK***	9270	4034	48414	0,075	2653	31830	0,049	1266	15191	0,024
EU-10	**35549**	**40155**	**481856**	**0,769**	**29171**	**350046**	**0,559**	**18855**	**226245**	**0,362**
Iceland****	25	257	3080	0,008	83	990	0,002	5	59	0,0002
Norway	351	1845	22145	0,031	898	10779	0,015	122	1464	0,002
Switzerland	454	3715	44585	0,062	2427	29124	0,041	1199	14391	0,020
EEA/EU-13	**36379**	**45972**	**551667**	**0,870**	**32578**	**390939**	**0,617**	**20181**	**242159**	**0,384**

Notes: *For Finland, predictions do not contain numbers for the Czech Republic, **For Italy and Spain the predictions contain numbers for Poland only, ***UK numbers do not contain Baltic countries: Estonia, Latvia and Lithuania. **** For Iceland the predictions contain numbers for Poland and Estonia only. Figure 5.IIIB in Appendix III shows prediction intervals for the mean gross migration rate across the three different convergence scenarios. Prediction errors for each pair of countries are available from the author upon request.

The highest predicted net flows from the new EU members are again seen for Germany with the average net migration flow being around 12,000 CEE migrants annually. This number is much lower than the one predicted in other studies, e.g. Fertig (2001), Fertig and Schmidt (2000).

Table 5.7 presents the development in migration stocks. Here, the annual net migration flow numbers are simply added to migration stock year by year. The total accumulated net increase in migration over the 12 years after the EU enlargement leads to a total number of 1.2 million migrants from those countries residing in the 13 EEA/EU countries in 2015. This is equivalent to 1.8% of the source countries' populations. Contrary to the predicted gross flows, the results in this study concerning the predicted stock are in the low end compared to other recent studies, but they get close to the findings of e.g. Boeri and Brücker (2001), Alvarez-Plata et al. (2003) and Dustmann et al. (2003) for Germany.

Table 5.7: Development in migration stocks from the new EU members in the 13 EEA/EU countries, 2004-2015.

Stocks	**Year 2000**		**High convergence scenario, year 2015**		**Medium convergence scenario, year 2015**		**Low convergence scenario, year 2015**	
Countries:		% of source		% of source		% of source		% of source
Belgium	9461	0,013	39536	0,055	26918	0,038	16139	0,023
Denmark	14903	0,021	54374	0,076	40006	0,056	26964	0,038
Finland	12849	0,018	34253	0,048	28642	0,040	25265	0,035
Germany	437104	0,611	637401	0,891	617372	0,863	597159	0,834
Greece	7157	0,01	37784	0,053	25182	0,035	14347	0,020
Italy	38400	0,054	95816	0,134	85726	0,120	75611	0,106
Netherlands	27897	0,039	85209	0,119	68505	0,096	52314	0,073
Spain	10441	0,021	38563	0,060	28723	0,059	21449	0,033
Sweden	75465	0,105	128072	0,179	110257	0,154	92677	0,129
UK	91597	0,142	150204	0,233	133224	0,207	116185	0,181
EU-10	**725274**	**1,034**	**1301212**	**1,847581**	**1164555**	**1,667**	**1038110**	**1,472**
Iceland	743	0,001	4746	0,007	2452	0,004	1343	0,002
Norway	9050	0,013	33218	0,046	21449	0,030	11731	0,016
Switzerland	13930	0,019	67582	0,094	51692	0,072	36526	0,051
EEA/EU-13	**748997**	**1,067**	**1406758**	**1,99504**	**1240148**	**1,772**	**1087710**	**1,541**

Notes: The prediction for Spain shows numbers for Poland and the Czech Republic only, the rest of the countries have missing migration records. UK numbers do not contain Baltic countries: Estonia, Latvia and Lithuania.

However, due to the fixed effects specifications those predictions assume no change in the immigration policy, i.e. no free movement. Obviously, the opening of the EU labor market is likely to affect the gross and net migration flows from the new EU countries.

The growth assumptions for the alternative high and low convergence scenarios lead to the average annual gross migration flows of 310,000 and 334,000, respectively. The accumulated gross migration flow over the predictive period 2004-2015 is then equal to 3.7 and 4 million migrants from the new EU members, for high and low convergence scenarios, respectively. This is equivalent to 5.4 and 5.8% of the source countries' populations, respectively. The magnitude of the range of estimated gross migration flows belongs to the higher end of the forecasts from the previous studies.

The results of predicted net migration flows under alternative scenarios show the average annual net increases in the stocks by 46,000 and 20,000 under assumptions for high and low convergence scenarios, respectively. This leads to a total number of 1.4 and 1.1 million migrants from those countries, which is equivalent to 2 and 1.5% of the source countries' populations residing in 13 EEA/EU countries in 2015. Such a development in the gross and net migration flows indicates that migration from the new EU member countries towards the "old" EEA/EU countries would remain a temporary phenomenon in the sense that besides the inflows there will be substantial outflows as well.

5.6. CONCLUSIONS

In this paper, I have made use of new sources of data that allow me to follow the actual migration from new EU member and EU candidate countries to the EEA/EU-13 countries over the period 1990-2000. Being able to observe migration behavior from these particular countries helps me to avoid problems related to (double) out-of-sample forecasts and to the assumption of invariance of migration behavior across a space that previous studies had to hold.

The results of the panel data analysis show indeed importance of controlling for pairs of countries unobserved heterogeneity. Further, some preliminary results regarding predictions of future gross and net migration flows suggest that the fears concerned with the large scale migration are hard to justify. Furthermore, these relatively modest migration flows from CEECs towards "old" EU members may in fact play a rather positive role in the destination countries that are facing a problem of declining and ageing populations.

Specifically, the average annual net increases in stocks from the 7 new EU member states are predicted to be between 20,000 and 46,000 depending on assumed convergence scenario. This leads to a total number of between 1.1 and 1.4 million migrants from those countries residing in the 13 EEA/EU countries in 2015, which is equivalent to 1.5–2% of the source countries' populations. Concerning the gross migration flows, the results show the total predicted accumulated gross flow over the period 2004-2015 being 5.4–5.8% of the source population, again depending on the assumed growth scenario. The results are relatively close to the findings of previous studies, although the predictions of gross migration potential belong to the larger ones, while the net migration potential belongs to the lower ones.

The results above show predicted net and gross migration flows and stocks in the case that the current immigration restrictions stay unchanged. But the opening of the EU labor market is likely to affect the migration flows from the new EU countries and thus especially in this case the Lucas critique can be applied. Therefore, it is an obvious task for future research to model and simulate the openings of the EEA/EU labor markets by relaxing the assumption of time-invariant employment rate in the EEA/EU destination countries. These are, however, steps to be carried out in future work.

REFERENCES

Alvarez-Plata, P., Brücker, H. and B. Sillverstovs (2003): "Potential Migration from Central and Eastern Europe into the EU-15 – An Update" *Report for the European Commission DG Employment and Social Affairs*, DIW Berlin.

Ammermueller, H., Heijke, H. and L. Woessmann (2005): "Schooling Quality in Eastern Europe: Educational Production during Transition". *Economics of Education Review,* Vol. 24, pp. 579-599.

Arellano, M. and S. Bond (1991): "Some Tests Of Specifications for Panel Data: Monte Carlo Evidence And An Application To Employment Equations" *Review of Economic Studies,* Vol.58, No.2, pp. 279 – 299.

Arellano, M and O. Bover (1995): "Another Look at the Instrumental Variable Estimation of Error Components Models" *Journal of Econometrics*, Vol. 68, pp. 29-51.

Baltagi H.B. (2005): "Econometric analysis of panel data". 3th ed. John Wiley & Sons, Inc. ISBN 0-470-01456-3.

Bauer, T. and K. Zimmermann (1999): "An Assessment of Possible Migration Pressure Following EU Enlargement to Central and Eastern Europe" *IZA Research Report No. 3.*

Belot, M. and S. Ederveen (2005): "Cultural and Institutional Barriers in Migration Between OECD Countries" *unpublished*.

Bentzen, J. and T. Engsted (1997): "Dynamic Modeling of Energy demand: A Guided Tour through the Jungle of Unit Roots and Cointegration". *Aarhus School of Business Working Paper 97-6.*

Boeri, T. and H. Brücker (2001): "Eastern Enlargement and EU-Labour-Markets: Perceptions, Challenges and Opportunities". *IZA Discussion Papers 256*.

Bruder, J. (2003): "East-West Migration in Europe, 2004 – 2015" *Thünen-Series of Applied Economic Theory Working Paper No. 40.*

Brücker, H. and B. Siliverstovs (2004): "The Macro-Determinants of International Migration in Europe: Evidence from Germany" Paper presented at the Conference "Immigration Issues in EU-Turkish Relations" Bogazici University.

Burda, M. (1995): "Migration and the Option Value of Waiting". *Economic and Social Review 27*, pp. 1-19.

Chiswick, B. (1996): "The Economic Consequences of Immigration: Application to the United States and Japan" in Weiner M. and T. Hanami, eds: Temporary Workers and Future Citizens? Japanese and U.S. Migration Policies. New York: New York University Press, pp. 177 - 208.

Chiswick, B. and T. Hatton (2002): “International Migration and the Integration of Labor Markets” *IZA Discussion Papers 559*.

Dustmann, C., Casanova M., Fertig, M., Preston, I. and C. M. Schmidt (2003): “The Impact of EU Enlargement on Migration Flows”. *Home Office Online Report 25/03.*

European Commission (2001): “The Economic Impact of Enlargement” Directorate General for Economic and Financial Affairs. ISSN 1608-9022.

http://europa.eu.int/comm/economy_finance/publications/enlargement_papers/2001/elp04en.pdf

Faini, R. and A. Venturini (1994): “Migration and Growth: The Experience of Southern Europe” *CEPR Discussion Papers No. 964*.

Fassmann, H. and Ch. Hintermann (1997): “Migrationspotential Ostmitteleuropa. Struktur und Motivation Potentieller Migranten aus Polen, der Slowakei, Tschechien und Ungarn”. *ISR- Research Report No.15*. Vienna.

Fertig, M. (2001): “The Economic Impact of EU-enlargement: Assessing the Migration Potential” Empirical Economics, pp. 707-720.

Fertig, M. and C.M. Schmidt (2000): “Aggregate-Level Migration Studies as a Tool for Forecasting Future Migration Streams” *IZA Discussion Paper 183*, IZA, Bonn

Filer, R. (1992): “The Impact of Immigrant Arrivals on Migratory Patterns of Native Workers” in Borjas G. and R. Freeman, eds.: Immigration and the Workforce: Economic Consequences for the United States and Source Areas. Chicago University Press, Chicago, pp. 93-134.

Granger, C.W.J. and P. Newbold (1974): “Spurious Regressions in Econometrics”. *Journal of Econometrics*, Vol.2, pp. 111 – 120.

Harris, J.R. and M.P. Todaro (1970): “Migration, Unemployment and Development: A Two-Sector Analysis” *American Economic Review,* 60(5), pp. 126-142.

Hatton, T. (1995): “A Model of UK Emigration, 1870-1913”. *The Review of Economics and Statistics*, pp 407-415.

Hatton, T. and J.G. Williamson (2002): "What Fundamentals Drive World Migration?," *NBER Working Paper No. 9159.*

Hille, W. and T. Straubhaar (2001): “The Impact of EU enlargement on Migration Movements and Economic Integration, Result and Recent Studies” In: OECD: Migration Policies and EU Enlargement. The Case of Central and Eastern Europe, OECD Paris.

Hunt, J. (1992): “The Impact of the 1962 Repatriates from Algeria on the French Labour Market”. *Industrial and Labor Relations Review, Vol. 45*, pp. 556-72.

Karemera D., Oguledo, V.I. and B. Davis (2000): “A Gravity Model Analysis of International Migration to North America” Applied Economics, Vol. 32 (13), pp. 1745-1755.

Lebhart, G. (2003): “Volkszählung 2001: Geburtsland and Staatsangehörigkeit” *Statichtische Nachrichter, Vol.4*.Vienna.

Madala, G.S. and S. Wu (1999): “A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test”. *Oxford Bulletin of Economics and Statistics*, pp.631-652.

Mayda, A.M. (2005): “International Migration: A Panel Data Analysis of Economic and Non-Economic Determinants” *IZA Discussion Paper 1590,* IZA, Bonn.

Pedersen, P.J., Pytlikova M. and N. Smith (2004): “Selection or Network Effects? Migration Flows into 27 OECD Countries, 1990-2000”, *IZA Discussion Paper, no. 1104,* IZA, Bonn.

Pedersen, P. J., Pytlikova M. and N. Smith (2005): “Migration into OECD countries 1990-2000”. In Parson and Smeeding (eds.): Immigration and the Transformation of Europe. Cambridge University Press. *In Print.*

Portes, J. and S. French (2005): “The Impact of Free Movement of Workers from Central and Eastern Europe on the UK Labour Market: Early Evidence”, *Department for Work and Pensions Working Paper No 18.*

Pytlikova M. (2005): “Where Did Central and Eastern Europeans Go and Why?” paper presented at the EALE/SOLE world conference, San Francisco 2005.

Roodman, D. (2005): xtabond2: Stata module to extend xtabond dynamic panel data estimator. Center for Global Development, Washington.

http://econpapers.repec.org/software/bocbocode/s435901.htm

Sinn, H.-W., Flaig, G., Werding, M. Munz, S., Düll, N. And H. Hofmann (2001): “EU-Erweiterung und Arbeitskräftemigration: Wege zu einer Schrittweisen Annäherung der Arbeitsmärkte“. Ifo study commissioned by the German Ministry of Labour and Social Order.

Sjaastad, L. A. (1962): „The Costs and Returns of Human Migration” *Journal of political Economy,* 70(5), pp. 80-93.

Wallace, C. (1998): “Migration Potential in Central and Eastern Europe”. *IOM Technical Cooperation Centre for Europe and Central Asia*. Geneva.

Zaiceva, A. (2004): “Implications of EU accession for international migration: and assessment of potential migration pressure” *CES IFO Working paper No. 1184*.

Zoubanov, N. (2004): “Assessing general and country-specific determinants of migration in the European Union. A panel data approach” paper presented at the IZA Summer School 2004.

APPENDIX I: *Previous literature*

Table 5.IA: Existing studies on CEE migration potentials.

Study	**Countries of origin**	**Destinations**	**Flows per year**	**Long-run (absolute)**
Hille, Straubhaar (2001)	10 CEECs	EU15	188.000-396.000	
Boeri, Brücker (2001)	10 CEECs	EU15	335.000 down to 100.000 in 2030	3.900.000 in 2030
Alvarez-Plata et al. (2003)	10 CEECs	EU15	335.000 down to 0 in 2030	3.817.409 in 2030
Bauer, Zimmermann (1999)	POL, CR, HUN, SR, SLO, ROM, BUL,	EU15	200.000 per year	883.505 – 2.614.544
Fertig (2001)	10 CEECs	Germany	67.101	1.409.119 in 2015
Fertig and Schmidt (2000)	POL, CR, EST, HUN	Germany	14.656 – 62.656/	293.122 – 1.253.000 accumulated inflows over 1998 – 2017
Sinn et al. (2001)	POL, CR, HUN, SK, ROM	Germany	240.000 – 270.000 down to 60.000-150.000 in 2020	
Dustmann et al. (2003)	AC-10	Germany/(UK)	20.459 – 209.651 (5.000 – 13.000) to 2010	
Zaiceva (2004)	10 CEECs	EU-15	300.000 – 500.000	3.500.000 – 5.000.000 over 2004-2014
Bruder (2003)	CEEC-8 without Malta and Cyprus	EU-15	238.063-273.300	1.700.000 (net) 3 000 000 (gross) 2004-2015

APPENDIX II: *Data*

5.IIA. List of countries included in the emigration flows' analysis:

Destination countries

Belgium, Denmark, Finland, Germany, Greece, Iceland, Italy, the Netherlands, Norway, Spain, Sweden, Switzerland and the UK

Source countries

The Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, the Slovak Republic

5.IIB. Description and definitions of the variables used in the paper and their source:

Gross flow of migrants from country i to country j

Source: National statistical offices and "Trends in International Migration" SOPEMI OECD.

Stock of foreigners from country i in country j

Source: National statistical offices and "Trends in International Migration" SOPEMI OECD.

Total population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship - except for refugees not permanently settled in the country of asylum, who are generally considered part of the population of their country of origin.

Source: World Bank.

GDP per capita (constant 1995 international $) based on purchasing power parity (PPP). PPP GDP is gross domestic product converted to international dollars using purchasing power parity rates. An international dollar has the same purchasing power over GDP as the U.S. dollar has in the United States. GDP at purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in constant 1995 international dollars.

Source: World Bank, International Comparison Programme database

Unemployment, total (% of total labour force): Unemployment refers to the share of the labour force that is without work but available for and seeking employment. Definitions of labour force and unemployment differ by country.

Source: World Bank: International Labour Organisation, Key Indicators of the Labour Market database.

Distance between countries – distance between capitals in km.

Source: MapInfo, own calculations.

5.IIC. Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
year	1092	1994.5	3.453634	1989	2000
Flows ij	717	2750.865	16253.94	0	260266
Stock ij	798	9565.734	34539.64	0	301366
POPj	1092	2.31e+07	2.51e+07	252700	8.21e+07
POPi	1092	1.03e+07	1.20e+07	1354666	3.87e+07
GDPpcPPPj	1092	21176.55	3930.378	12383.73	32227.57
GDPpcPPPi	1053	8561.24	1946.647	4963.767	12839.68
Uj	1092	7.907412	4.513214	.5008767	23.66247
Ui	962	9.703817	4.914661	.3	18.9
Dist ij	1092	1347.302	664.5027	86.5	3070

5.IID. Testing time-series properties of the data:

The panel structure of the data enables me to analyze migration behavior in a two-dimensional framework, across space and time. Therefore, I look at the time-series properties of my data in the following part. Specifically, I test for the existence of unit roots in the variables. This is quite important, as if a series turns out to have a unit root, then its variance is infinite, therefore standard statistical inference is not valid asymptotically (Granger and Newbold, 1974). On the other hand, if a long-run relationship between the dependent variable and explanatory variables exists, i.e. if the variables are cointegrated, the outcome may be different as nonstationarity in this case may provide a possibility for obtaining super-consistent results, see Bentzen and Engsted (1997) for a discussion on these issues.

Although unit root's tests are commonly used in time-series data analysis[114], their panel data applications are not that developed.[115] Till now, the following tests are available: Breitung and Meyer (1994) test, Im, Pesaran and Shin (IPS) test and the Fisher test (Maddala and Wu, 1999). However, all tests except the last mentioned one, require balanced panels, which is not a case for my dataset. Therefore, I decided to test for unit roots by using the Fisher test. The test combines p-values from N independent unit root tests, as developed by Maddala and Wu (1999), and

[114] A standard procedure is to run a Dickey and Fuller test.

[115] A good overview of panel data unit root tests in the context of migration is given in Zoubanov (2004).

computes the statistic $\lambda = -2\sum_{i=1}^{N} \ln p_i$, where p_i is observed significance level (p-value) in the ith individual unit root test. Based on the p-values of individual unit root tests, Fisher's test assumes that all series are non-stationary under the null hypothesis against the alternative that at least one series in the panel is stationary.

The results of the Fisher's unit root tests for migration flow, stock and other variables, which enter the model, are given in Table 5.IIA. For migration flows and stock the Fisher test tends to reject the hypothesis that all flows and stock series in the panel are non-stationary.[116] In the case of GDP per capital in PPP and employment rates, the Fisher test cannot reject the Ho that all series are non-stationary, thus these variables have unit roots. Their first differences reveal that they are integrated of order one.

Table 5.IIA: Fisher's unit root test (ADF) for the entire panel of destination and source countries.

Variables	**in levels**		**in differences**	
	Chi-squared	*p-value*	*Chi-squared*	*p-value*
Immigration Flows $\ln(m_{ij})$	chi2(320) = 906.980	0.0000	chi2(492) = 2116.374	0.0000
Immigration Stock $\ln(s_{ij})$	chi2(240) = 1825.456	0.0000	chi2(384) = 2244.920	0.0000
GDP per capita PPP $\ln(GDP_j)$	chi2(810) = 136.058	1.0000	chi2(810) = 1440.443	0.0000
GDP per capita PPP $\ln(GDP_i)$	chi2(810) = 654.860	1.0000	chi2(810) = 912.191	0.0070
Employment rate $\ln(e_j)$	chi2(360) = 139.688	1.0000	chi2(310) = 1293.474	0.0000
Employment rate $\ln(e_i)$	chi2(350) = 126.134	1.0000	chi2(252) = 812.291	0.0000

In order to look into the stationarity properties of flow and stock variables, I run the Fisher test for flows and stock variables for each destination country separately, see Table 5.IIB. The test shows that the results are mixed for different destination countries with half of the destination countries having stationary flows and stock and half having a unit root. This means that there exist specific co-integration relations for particular countries or pair of countries. Given the short panel of the data, one

[116] The null hypothesis is that all time-series in the panel are stationary I(0) – if the hypothesis is rejected, it does not imply that the hypothesis is rejected for all groups, i.e. there may be a mixture of I (0) and I(1) variables in the panel.

should be careful with any strong interpretation from the panel unit root tests. But it can at least give a feeling that the stationarity/non-stationarity in variables may be different for different countries.

Some previous studies, e.g. Alvarez-Plata et al. (2003), Brücker and Siliverstovs (2004), argue that there is no equilibrium between migration flows and other explanatory variables on the basis of unit-root tests. They argue that such a relationship exists between stocks and explanatory variables only. However, the panel unit root tests on my data reveal that the results may be different for different countries.

Table 5.IIB: Fisher's unit root test (ADF) for each particular destination country.

	Immigration Flows		**Immigration Stock**			**Immigration Flows**		**Immigration Stock**	
Austria	chi2(28) = 148.0791	0.0000	-	-	**Italy**	chi2(10) = 15.1855	0.1254	chi2(4) = 0.012	1.000
Belgium	chi2(22) = 6.0195	0.9997	chi2(20) = 151.197	0.0000	**Netherlands**	chi2(28) = 68.7396	0.0000	chi2(28) = 52.521	0.0001
Denmark	chi2(28) = 50.6672	0.0054	chi2(20) = 249.530	0.0000	**Norway**	chi2(20) = 72.1684	0.0000	chi2(16) = 118.659	0.0000
Finland	chi2(28) = 53.6129	0.0025	chi2(20) = 460.429	0.0000	**Spain**	chi2(14) = 0.0276	1.0000	chi2(6) = 0.491	0.9987
France	chi2(28) = 45.2611	0.0208	-	-	**Sweden**	chi2(28) = 117.4586	0.0000	chi2(20) = 322.4328	0.0000
Germany	chi2(26) = 18.5743	0.8539	chi2(20) = 123.195	0.0000	**Switzerland**	chi2(28) = 154.4900	0.0000	chi2(20) = 236.427	0.0000
Greece	chi2(28) = 121.2409	0.0000	chi2(20) = 28.878	0.0902	**United Kingdom**	chi2(12) = 4.1617	0.9803	chi2(14) = 25.193	0.0327
Iceland	chi2(22) = 30.2483	0.1126	chi2(20) = 19.445	0.493					

APPENDIX III: *Standard errors for the predictions*

Figure 5.IIIA: Development of the observed and predicted mean gross migration rate and prediction intervals for the mean gross migration rate; 1989-2015.

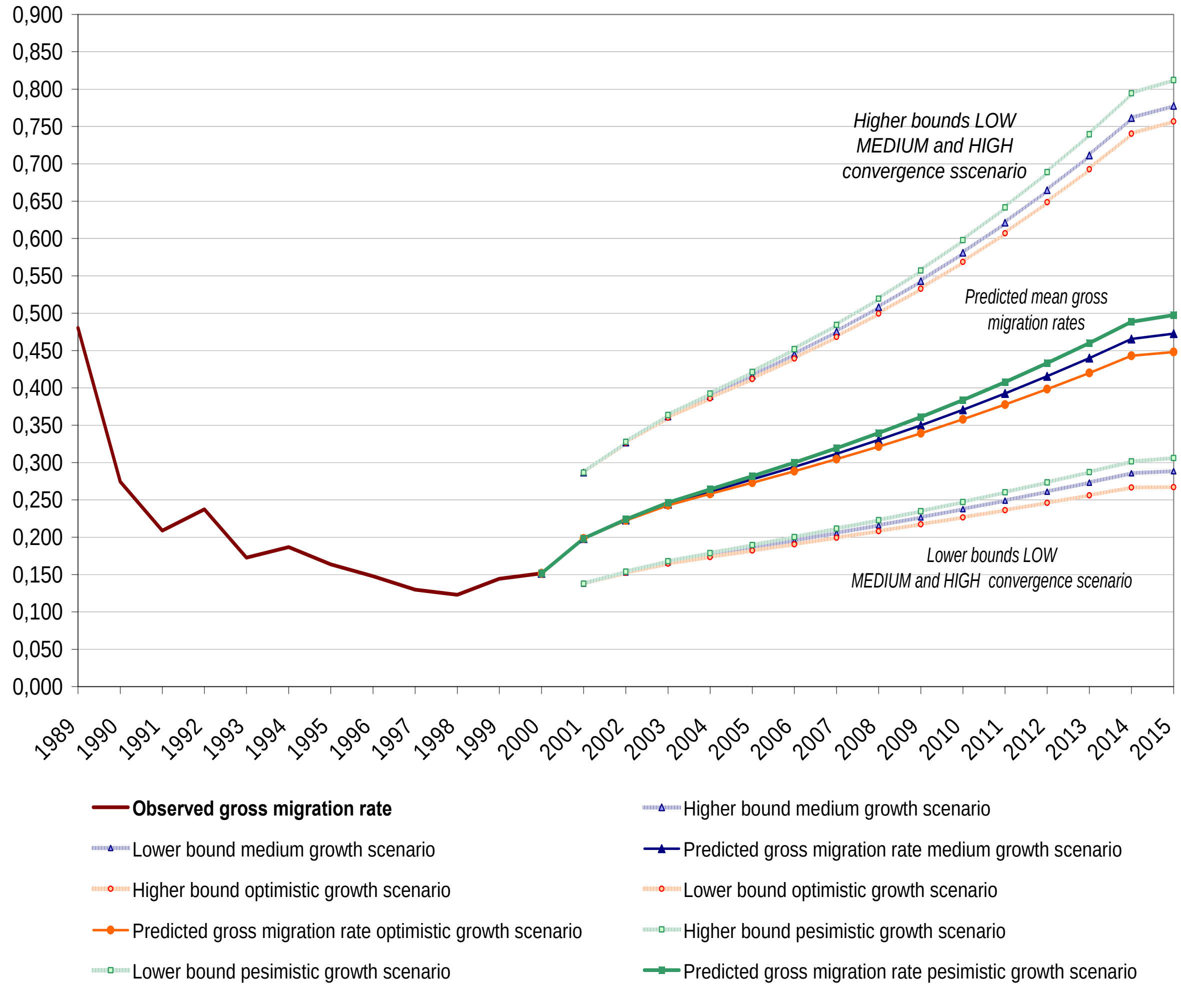


Figure 5.IIIB: Development of the observed and predicted mean net migration rate and prediction intervals for the mean net migration rate; 1990-2015.