EERI
Economics and Econometrics Research Institute

Can we use NEG models to predict migration flows? An example of CEE accession countries

d'Artis Kancs

EERI Research Paper Series No 1/2005

EERI
Economics and Econometrics Research Institute
Avenue de Beaulieu
1160 Brussels
Belgium

Tel: +322 299 3523
Fax: +322 299 3523
www.eeri.eu

Can we use NEG models to predict migration flows? An example of CEE accession countries*

d'Artis Kancs†

Abstract

In this paper we develop an analytically solvable and structurally estimable economic geography model and apply it to predict migration flows for the period following the CEE's integration with the EU. The main innovation of our approach is that it endogenises both explanatory variables and the migration rate. The model's parameters are estimated econometrically using a migration equation, which is derived entirely from the theoretical NEG model.

Our simulations show that even relatively moderate changes in some of the explanatory variables (such as transport costs) can actuate unpredictable changes (both in sign and magnitude) in other explanatory variables (such as wages). Keeping these explanatory variables fixed, as in reduced-form models, would produce biased results. Our empirical findings advocate that there is enough evidence to predict a selective migration among the three Baltic states. However, labour mobility in the Baltic countries is sufficiently low to make the swift emergence of a core-periphery pattern very unlikely at this geographical level.

Keywords: Migration, economic geography, European regions, integration.
JEL classification: F12, L11, R12, R23.

*The author acknowledges helpful comments from Paul Cheshire, Matthieu Crozet, Jurgen Essletzbichler, Julda Kielyte, Geoff Meen and Henry Overman as well as seminar participants at LSE and CEPR. The author acknowledges financial support from the Research Grant 'International Labor Migration in a Globalizing Economy'. The author is solely responsible for the content of the paper. The views expressed are purely those of the author and may not in any circumstances be regarded as stating an official position of the European Commission.

†London School of Economics, Economics and Econometrics Research Institute (EERI), European Commission - Joint Research Centre (IPTS), Catholic University of Leuven (LICOS). E-mail: kancs@lse.ac.uk.

1 Introduction

Freedom of movement is one of the fundamental principles upon which the European Union (Community) was once founded. The free movement of workers within the Single European market is an integral part of the Treaty of the European Community. It belongs to the *acquis communautaire* that has to be granted reciprocally to citizens from old and new EU Member States (*European Commission* 2004)[1]. The recent and unprecedented EU enlargement, however, attracted a huge public attention and remains a highly controversial issue at all political and societal levels.

The high relevance of the topic has generated a large body of theoretical and empirical literature that attempts to predict the size of possible labour migration in the years following the accession. Most of these studies are based on reduced-form econometric models and, in addition. In this paper we argue that the reduced-form approach, especially when applied to economies in transition, where explanatory variables are to unpredictable changes, is biased (Section 3.2). The current paper develops an alternative methodology - a structurally estimable NEG model of migration flows and provides an empirical application of the model (Sections 2.1 and 2.2).

We proceed in three steps. In a first step we derive a tractable migration equation from a new economic geography model, where migration across regions eliminates real wage differentials. The canonical economic geography model we use represents an analytically solvable version of *Krugman's* (1991) core-periphery model. In a second step we use data on historical migration experience to estimate coefficients of the derived migration function. The estimated coefficients provide estimates of key parameters of the NEG model. In a last step we use the NEG model and the estimated parameters for simulations to a period after the CEE integration with the EU, when free movement of workers is introduced. European integration is modelled by altering the model's parameters - reducing inter-regional transport costs.

The rest of the paper is structured as follows. Section 2 gives an overview about historical migration patterns and institutional settings in the new EU member states. After providing an outlook about volume and direction of potential migration flows, Section 3 critically assesses the existing literature and discusses methodological shortcomings of previous migration studies. Section 4 formally presents the canonical economic geography model and derives an estimable migration equation from the theoretical NEG model. The migration function is estimated and estimation results are presented in Section 5. The NEG model's simulation results and policy recommendations are presented in Section 6. Section 7 concludes.

[1] In this paper EU-15 are referred as old EU Member States and CEE-10 accession countries as new EU Member States.

2 Migration in CEE

This Section focuses on historical labour movements as well as more recent patterns of international migration concerning the new EU members, in particular the CEE accession countries. It also identifies migration policies, which will determine migration flows after the CEE integration with the EU, when in selected EU member states free movement of workers is introduced. In particular, the relationship between migration and the economic and policy factors in old and new EU member states is discussed. The Section concludes by identifying potential origin and destination countries as well as the size of the future migration flows in Europe.

2.1 Historical overview

In this Section we argue that because of a relatively low international workers' mobility in the new member states when compared to the old EU member countries, inter-regional migration within the CEE will be at the heart of matters after the EU enlargement and not the often discussed East-to-West migration. This is best shown by exploring historical patterns of international migration. In particular, to provide a background for the analysis of current migration trends, developments of the population movements from the past should be studied, most importantly from the period after 1945, bearing in mind the unique character of mass migratory flows in Europe in the direct aftermath of the second World War. Therefore, we start our analysis with the Soviet period, when the CEE accession countries were under the Socialist regime.

After the turmoil of post-war migration, which was the consequence of new post-Yalta World and the European order, the two parts of Europe lived their own lives. Western Europe was a booming economy, with an inelastic labour market, which was supported by imported foreign labour force. Only the oil crisis of 1973 resulted in a change in migration policies, but at that time the momentum of immigration was high and Western Europe was becoming multiethnic and multi-cultural (*Kielyte & Kancs* 2002).

In contrast, in Central and Eastern Europe the population migration was deemed to be a political issue and, therefore, controlled by the government and by the Communist party. One of the common features of migration movements in the former socialist countries was the East-to-West direction of most of the long-term population flows, very few return migrations, and hardly any international migration beyond the former Soviet bloc[2]. Baltic states, due to the fact they were part of the former Soviet Union, had a completely different migration experience (*OECD* 2003).

The three Baltic states (Estonia, Latvia and Lithuania), with their shared history of constituent parts of the Soviet Union from 1940 until 1991, were characterised by

[2]The exception were temporary workers hired on the basis of intergovernmental agreements between the socialist countries.

centrally planned migration patterns. Although international migration as such was hardly existent due to a very strict external movement control and a rigid visa policy, there were significant population movements between the republics of the former USSR. In the communist period, Estonia, Latvia and Lithuania observed strongly positive migration balance (Figure 1), comprised mainly of immigrating Russians, many of whom were the Soviet military personnel. This migration phenomena contributed to significant changes of the ethnic structures in all three Baltic republics (*Kielyte* & *Kancs* 2002).

After more than 30 years, net migration became negative in all three Baltic countries for the first time in 1990 (Figure 1)[3]. Figure 1 reports that immigration flows, having begun to decline in the late 1980s fell sharply in the early 1990s since when they have stabilised at a historically low level. The net emigration was increasing in the case of Lithuania until 1992 and in the case of Estonia and Latvia until 1994. The negative total net migration was to a large extent caused by international out-migration of the so-called 'Russian-speaking population'.

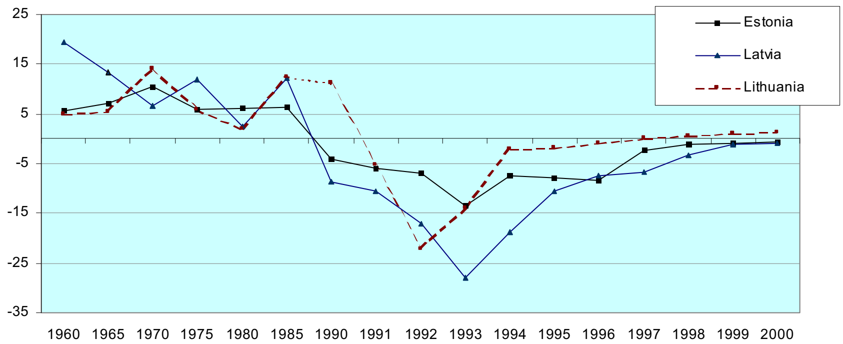

Figure 1: Net Migration in the Baltic States in Thousands, 1960-2000

Data for 2000s indicate the most recent migration trends among the three Baltic states. In 2000, almost 2,913 immigrants were registered in Latvia, just over 2,536 in Lithuania, and almost 1,156 in Estonia. Since 1997, only Lithuania's net migration balance became positive again and amounted to 1,100 persons in 1999. In 2005 the international migration balance (not in the figure) remains negative in Latvia (1,800) and in Estonia (600).

Bearing in mind that a substantial bulk of migration, which happened during and after the fall of Communism in the 1960s till late 1980s and early 1990s, is unlikely to occur again, we may conclude that international mobility of population in the three Baltic countries is rather low. The relatively low historical peoples' mobility in the three Baltic states when compared to 'old' EU member countries suggests

[3] Positive numbers stand for immigration and negative for emigration.

that workers migration within the same geographical region might be at the heart of matters after the EU enlargement.

2.2 Migration policy

In this Section we argue that the potential for an inter-regional migration among different parts of the CEE accession countries, which do not restrict access to their labour markets, is much higher than the potential for an East-West migration, where CEE accession countries' citizens are excluded from the free movement rights. The basic rights attached to citizens of the European Union are set out in part two of the EC Treaty (Art 17 – 22). These include the right to travel, work and live freely in another country, the right to vote and stand in municipal elections in one's country of residence, and the right to diplomatic and consular protection.

As from 1st of May these rights should apply to citizens throughout EU-25. However, the extension of free movement rights to the additional 75 million new EU citizens turned out to be a particularly sensitive and 'popular' topic marked by many intensive debates during the accession negotiations (*European Commission* 2004). The trade-off was given by the ambitions and pressure exerted by the new member states on the one hand and by the anxiety about possible negative effects on the labour market and employment conditions voiced by some of the old member states on the other hand. The solution was found in a rather complex 2 years + 3 years + 2 years transitional arrangement referring to workers and proposed by the Commission. In the process of negotiations the transitional measures have been agreed with all Eastern and Central European new member states and were included in the Accession Treaty signed on 16 April 2003[4].

Table 1. EU member states with open labour markets

Country	Measures (years)	Country	Measures (years)
Cyprus	–	Malta	–
Czech Rep.	–	Poland	$2(+3+2)^*$
Estonia	–	Slovakia	–
Hungary	$2(+3+2)^*$	Slovenia	–
Ireland	–	Sweden	–
Latvia	–	UK	Registration[†]
Lithuania	–		

[*]on a reciprocal basis, does not apply to CEE; [†]limited access to welfare benefits. Source: *European Commission* (2004)

The Accession Treaty itself does not clearly specify what form the transitional national measures may or must not take. The only guideline provided by this act is

[4]The possibility of derogation from the free movement of workers principle is set out in Annexes V and VI, VII – X and XII – XIV6 attached to the Act on Accession.

given in Paragraphs 13 and 14 stipulating that national measures applied may not be more stringent than those were applicable at the time, 16 April 2003. Therefore, member states resorted to a variety of different restrictions ranging from limitations depending on sector or type of work, quota arrangements, to work permits granted only when a national cannot be found to fill the vacancy. As shown in the following, all these measures are also to be applied in different timescales. We start our assessment with those EU member states, which do not restrict access to their labour markets.

According to Table 1, most of the EU member states (13 out of 25) do not restrict access to their labour markets in the enlarged European Union. Ireland, the UK, and Sweden are the only three old EU member states that decided to open up their labour markets as from the first day of the EU enlargement (Table 1). *None of the new member states have requested the Commission's authorisation to restrict access to their labour markets by new EU member states nationals.* Only two Central and Eastern European accession countries, Poland and Hungary, are known to be applying reciprocal measures. However, the labour market measures in Poland and Hungary are applied only with reference to nationals from the old member states (Table 1).

In contrast, a vast majority of the old EU member states set up adequate measures providing for the application of different work permit schemes for workers from the Central and Eastern European accession countries[5] (Table 2).

Table 2. Countries with restrictive labour markets

Country	Transitional measures (years)
Austria	2 (+3 +2 planned)
Belgium	2 (+3 +2 possible)*
Denmark	2 (+3 +2 possible)*
Finland	2 (+3 +2 possible)*
France	2 (+3 possible)*
Germany	2 (+3 +2 planned)
Greece	2 (+3 +2 planned)
Italy	2 (+3 possible)*†
Luxembourg	2 (+3 +2 possible)
Netherlands	2 (+3 +2 possible)*†
Portugal	2 (+3 +2 possible)*†
Spain	2 (+3 +2 possible)

*Current work permit system remains; †immigration quota (ca. 20,000 p.a.). Source: *European Commission* (2004)

For example *Belgium* retains its pre-enlargement work permit system with permits

[5] The Declarations annexed to the Final Act of the Accession Treaty have not been binding on any of the old member states. Contagious as a virus, the scare of an 'influx of migrants from Central and Eastern Europe' led some of the member states to 'rethink' their promises and as the 1st of May 2004 drew closer, in many cases to apply or announce more protectionist measures.

A (for all salaried workers) and B (for temporary employment) for a minimum period of two years. Workers wishing to take up employment in the *Netherlands* are also required to obtain a work permit, even though the government has abandoned its earlier intention of introducing quotas. In a number of sectors granting of such a permit is subject to simplified procedures, where the waiting time does not exceed two weeks. Nevertheless, work permits for all other jobs, falling outside the scope of specified sector 'relaxations' are granted only when a Dutch national (or national of other old Member State) willing to take the vacant post cannot be found. The situation is similar in *Finland*, where under national law, which is applicable to nearly all new EU nationals for a minimum period of two years, work permit will only be granted provided that the vacancy cannot be filled by a Finnish worker (Table 2). The system applied by *France* 'prima facie' may seem very similar to that operated by the Dutch, where current work permit policy applicable to salaried workers has been maintained. France foresees also possibilities of opening the labour market in specific professional sectors and currently the work permit requirement does not apply to students and researchers. Nevertheless, according to announcements made by this country's representatives, the system will be in place for a period of 5 years (Table 2).

Yet stricter work permit scheme is present in *Denmark*, *Italy* and *Portugal*. *Denmark* applies a system, under which citizens of the CEE accession countries are allowed to obtain a work permit only once they obtain an official residence permit and only for full time employment. *Italy* operates a work permit scheme, which is automatically limited by a fixed quota of 20,000 workers coming from the CEE accession countries in the year 2004. Only in cases of certain sector specific professions, work permits will be issued outside the scope of the quota fixed for 2004. Similarly in *Portugal*, for the period of two years after enlargement the current system of work permits granted within quotas set every two years (covering all foreign nationals with the exception of EU-15) will be maintained (Table 2).

Austria and *Germany*, which have traditionally been the two countries receiving the vast majority of migrants from Central and Eastern Europe, voiced their concerns about the probable negative impact of migration on employment markets most loudly. Therefore, both continue applying national restrictions (i.e. work permits schemes) and provisions arising from bilateral agreements signed between themselves and individual new Member States. Both countries are also allowed to apply certain restrictions on freedom to provide services (Table 2).

We may conclude that, with few exceptions, the CEE accession countries' citizens are virtually excluded from the free movement rights at least until the end of the decade. Although, as shown in the following Section, migration to the old EU member states is an intensively researched topic, this Section sheds doubts about the usefulness of studying impacts of East-to-West migration ignoring the potential East-to-East migration. Indeed, introduction of free movement of workers among the new EU member states might actuate migration within CEE, as all legal barriers to labour movement have been abolished and inter-regional welfare differences within CEE are increasing rapidly. Bearing in mind that per capita income in some NUTS II regions is twice as high as in other regions in the CEE accession countries, taking into account language and cultural similarities, and recognising that since enlargement there are no legal barriers to workers movement within CEE, it is reasonable to assume a high migration potential between low wage regions and high wage regions within the new EU member states.

3 Existing literature

The high relevance of migration topic in the enlarged EU has generated a large body of empirical literature that attempts to predict the size of possible labour migration in the years following the accession. In this Section, the most influential work is critically reviewed. We introduce main past achievements in the field and show theoretical and empirical gaps in the existing literature.

3.1 Empirical limitations

Studies that forecast future migration flows from the CEE accession countries to the EU are typically based on econometric models that make predictions using historical data on migration flows (see *Fertig* 2000, *Boeri* & *Brücker* 2001, *Bauer* & *Zimmermann* 1999). A variety of estimation specifications are used in this literature. Usually, the theoretical bases for the empirical specification are simple economic arguments that relate migration to differences in returns to human capital and costs of migration.

One of the basic assumptions fundamental to all models is the free movement of workers having already replaced the current regime of transitional measures. Hence none of the temporary regulations and administrative restrictions are considered in these studies. However, as shown in Section 2.2, migration is heavily regulated in the old EU member states. While in the trade literature it would be impossible to make any predictions about the trade flows without considering import tariffs and taxes, the ignorance of existing policy measures seems to be a standard practice in migration literature. This calls for a modelling strategy rather than comparing the two corner solutions - autarky versus free mobility. The framework developed in the current paper allows us to investigate migration responses to gradual shifts in migration barriers.

A common way of estimating the migration potential from CEE accession countries is through extrapolation exercises, which take migration flows from Southern Europe to the West and North European countries in the 1950s and 1960s as point of reference (*Bauer* & *Zimmermann* 1999). However, as shown in Section 2.1, there are important differences between the conditions of South-to-North migration and migration from the CEE accession countries. First, in Western Europe labour markets were characterised by full employment and shortages of manual workers in the main receiving countries (Belgium, France, Germany and Switzerland) until the first oil price shock in 1973. Today, on the opposite, unemployment rates are rather high in the main receiving countries in the EU. Second, the transition process in some of the CEE countries is not yet complete, so rates of structural change and job turnover are higher in the CEE accession countries than in traditional sending countries. These remarkable differences in macro-economic environment suggest that the extrapolation of the South-to-North migration experience is reliable in relative terms, but can be hardly used to estimate migration levels. The data set compiled in this paper is unique in sense that it allows us to estimate bahavioral parameters using past migration experience from the same countries we will make prediction.

At the stage of empirical implementation, usually there are additional assumptions imposed that are hard to justify. For instance, the assumption made in some studies (e.g. *Fertig* 2000, *Boeri & Brücker* 2001) that the slope (response) parameters are the same for all countries is very strong. It implies that immigration from countries such as Rumania, Bulgaria responds to a change in relative GDP in the same way as immigration from Slovenia. As a consequence, the coefficients can be biased and the migration behaviour may deviate from that in the sample on which the estimates are based. The accuracy of these calculations is, therefore, of the status of a rough guess rather than that of a rough estimate. In this study we estimate bahavioral parameters separately for each CEE accession country, which allows to obtain and use in predictions country-specific migration elasticities.

3.2 Methodological issues

This Section shows that despite their popularity, reduced-form models have quite a few limiting assumptions, which are hard to justify. Most of the enlargement papers (e.g. *Boeri & Brücker* 2001, and *Fertig* 2000) refer to *Hatton's* (1995) more elaborate model as a motivation for their econometric specification. Therefore, we discuss *Hatton's* model in a more detail. *Hatton* develops his model assuming that the individual migration decision is determined by considerations about relative earnings, employment and non-pecuniary costs of migrating to a destination country. His estimation equation is a relationship between migration rates and differences in key economic indicators, like income and employment. In addition, the costs of migration enter the formulation proxied by the stock of immigrants from the individual's origin country. Although carefully derived, *Hatton's* model is based on a number of ad hoc assumptions. He clearly acknowledges this by saying that '*it is worth emphasising that the model developed here is only one among many different specifications that could be developed*'.

First of all, the explanatory variables (such as per capita incomes, unemployment rates etc. both in the receiving and sending countries) are kept exogenous from the estimated model. Reduced-form models of fitting a relatively saturated specification to the observed migration data, typically including substantial number of economic variables on the right-hand side of the regression, in order to assess migration potential and predict future migration flows, require either strong assumptions of temporal stability of the behavioral relationships to hold or one has to have a precise notion about the development of these conditioning variables in the future. Unfortunately, and in contrast to developed economies, it s notoriously difficult to predict economic variables in transition economies undergoing structural changes. Therefore, reduced-form models with fixed explanatory variables are hardly justifiable for predicting migration flows in the CEE transition economies. The framework developed in this paper can make make future predictions by allowing both left and right hand site variables to adjust to changes in economic environment.

Last but not least are problems related with the endogeneity of these regressions. It is widely acknowledged that economically driven migration is largely determined by differences in per capita income, unemployment rate, labour market regulations and social policies among host and destination countries, as well as by moving costs. *Krugman* (1991) has shown that the interaction of labour migration across regions with increasing returns and trade costs creates a tendency for firms and workers to cluster together as regions integrate. Because of the *price-index effect* and *demand & cost linkages* this leads to a lower manufacturing price index, lower productions costs, higher productivity and higher real wages. Thus, migration rate depends on wages, income, price index and many other variables, which traditionally are used as regressors. These variables, however, turn out to be a function of a migration rate. Because of this reverse causality, reduced-form models' estimates will likely be biased upwards. The general equilibrium nature of our model accounts for both the endogeneity and the reverse causality problems.

4 Theoretical framework

Different parts of economic theory are well aware of the above discussed problems and have proposed solutions to cope with them. Two of the fields seem to be particularly relevant in the context of this study - the economic geography and the structural econometrics. Therefore, in order to overcome shortcomings of the reduced-form models, we propose a framework which combines the two approaches by taking advantages of each of them. For this purpose we take a canonical economic geography model derived from *Forlsid* & *Ottaviano* (2003), which in turn represents an analytically solvable version of *Krugman's* (1991) core-periphery model. In that model, we are able to study impacts of increasing regional integration on the location decisions of workers and firms.

By adopting the NEG model as a theoretical framework, we are able to cope with most of the limitations explained above. While most reduced-form models in the literature assume that a long-run equilibrium relationship between migration flows and the explanatory variables exists, an implicit assumption of the NEG model is that an equilibrium between migration stocks and the explanatory variables emerges in the long-run. While the flow-model is based on the assumption that all individuals are homogeneous with regard to their preferences and human characteristics, the NEG model assumes that individuals are heterogeneous. As a consequence, for a given difference in expected utility levels, the equilibrium migration stock is achieved eventually when the benefits of migration equals its costs for the marginal migrant. Note that the NEG model does not rely on the assumption that all migration is permanent. In contrast, under the assumption of heterogeneity, the duration of migration varies across individuals. In equilibrium, the emigration from some cohorts equals the return migration from some other cohorts, as long as the rate of natural population growth is similar in the home population and the migrant population. Econometric estimations are required to satisfy properties (ii) and (iii).

4.1 The NEG model: a non-technical overview

This Section spells out the R-region version of *Forlsid* & *Ottaviano* (2003) economic geography model. The 'world' consists of R regions, endowed with two factors, an immobile factor (L) and a mobile labour (H). Both factors are used in the production process. Regional supplies of the immobile factor are fixed: each region contains L_r units of the immobile factor. Labour, however, is inter-regionally mobile. The world hosts H units of labour: $H = H_1 + H_2 + ... + H_r + ... + H_R$ with $r \in \{1,..,d,..,o,..,r,..,R\}$. Workers migrate among regions maximising their utility and at the end of each period workers are endogenously distributed among regions (H_r stands for regions' initial endowment with labour, and $\hat{H}_r$ - for regions' endowment with labour at the end of a period (after migration)).

There are two sectors: traditional sector (A) and manufacturing (X). All goods are traded among all regions. 'Traditional' sector is perfectly competitive and immobile sector, and serves as a numeraire. The traditional good is homogenous and produced under perfect competition. It is assumed to be traded at zero cost, both inter-regionally and internationally. Therefore, its price equalises everywhere: $p_{A1} = p_{A2} = ... = p_{Ar}... = p_{AR}$. We chose units such that $p_A = r_A$ in each region. Recognising that the traditional good is a numeraire, this implies that $p_{Ar} = 1, \forall r$.

Monopolistically competitive 'manufacturing' produces a differentiated good and stands for all increasing-returns and mobile production activities in the economy. Product markets of all R regions are separated by trade costs. Manufacturing varieties produced in a region r are sold by firms at mill price, and the entire transaction cost is borne by consumers. Trade costs are of '*iceberg*' type: when one unit is shipped, priced p, only $\frac{1}{T}$ actually arrives at its destination. Therefore, in order for one unit to arrive, T units have to be shipped, increasing the price of the unit received to pT. Cross-border trade of manufactured goods is subject to such trade costs, which differ across regions. We also assume that trade costs are symmetric for any pair of regions, i.e. $T_{od} = T_{do}$, where o is the region of origin and d is the destination region, and

$r \in \{1, .., d, .., o, .., r, .., R\}$.[6]

4.2 Consumption

All consumers share the same quasi-linear utility function:

$$U = \alpha \ln C_x + C_A \qquad \text{with} \qquad \alpha > 0 \tag{1}$$

C_x is a composite index of the consumption of the manufactured good, C_A denotes consumption of the traditional good. The composite index C_x is defined by the following CES function:

$$C_x = \left[\sum_{j=1}^{N} x_j^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{2}$$

where x_j represents consumption of a variety j of the manufactured good, N is the number of available varieties in the economy, and σ is the elasticity of substitution between two varieties ($\sigma > 1$). Given income Y, each consumer maximises his utility subject to the budget constraint $Y = C_A p_A + C_x p_x$, where $p_x = \left(\sum_j p_j^{1-\sigma} \right) \frac{1}{1-\sigma}$. Using (1) and (2), we can derive the following demand function, representing demand emanating from consumers of region d, addressed to a producer j located in region o:

$$x_{j,od} = p_{j,od}^{-\sigma} \frac{\alpha}{\sum_j p_{j,od}^{1-\sigma}} \tag{3}$$

Equation (3) contains the spatial framework. Each region produces N_r varieties of the manufacturing good. Iceberg trade costs imply that the price of each variety j produced in region o and sold in region d contains the mill price and the trade cost: $p_{j,od} = p_o T_{od}$ (because of the symmetry of all varieties produced in the same region, we henceforth omit the variety subscript j). We use T_{od} as a general expression, assuming that the trade cost between two regions is identical for both directions of trade flows, and that $T_{rr} = 1$. Using (2) and (3) we are able to derive the following industrial price index for each region d:

$$P_d = \left[\sum_{o=1}^{R} N_o \left(p_o T_{od} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{4}$$

Individual demand (3) can now be written as:

$$x_{od} = \frac{\alpha \left(p_o T_{od} \right)^{-\sigma}}{P_d^{1-\sigma}} \tag{5}$$

[6] The later assumption of symmetric trade costs indeed corresponds to the observed data in countries under analysis.

4.3 Production

Manufactured goods are produced in a monopolistically competitive industry that employs both the immobile factor and labour. The marginal cost in terms of immobile factor is unitary. Each producer has the same production function. Recalling that the immobile factor is rented at a rent that is set equal to one, the total cost of producing x_j units of variety j in region r is $TC_r(x_j) = W_r H_j + L_r x_j$, where W_r represents the compensation of labour in region r. Hence, $TC_r(x_j)$ contains a fixed cost that corresponds to one unit of labour input, i.e. $H_j = 1$ and marginal cost in terms of the immobile factor. The fixed cost gives rise to increasing returns to scale.

As usual in a monopolistic competition framework, we assume that there are a large number of manufacture firms, each producing a single product. Hence we obtain the constant mark-up equation for profit maximising firms:

$$p_o = \left[\frac{\sigma}{\sigma - 1}\right], \quad \forall r \tag{6}$$

where p_o is the price of a variety produced in o. The restriction $\sigma > 1$ ensures that p_o is always positive.

The equilibrium output of a firm producing in region o is given by market clearing for each variety. Using (5), we can express output as:

$$X_o = \sum_{d=1}^{R} (H_d + L_d) T_{od} x_{od} \tag{7}$$

and the profit function of a representative firm located in r is:

$$\Pi_r = p_r X_r - X_r - W_r \tag{8}$$

The number of varieties produced equals the number of firms located in that region, which is linked one to one to the number of workers. Thus, $H_r = N_r$. The zero-profit condition in equilibrium implies W_r adjustment. Using (6) and (8), we obtain:

$$X_r = W_r (\sigma - 1) \tag{9}$$

4.4 Equilibrium

In a short run, workers are immobile between regions, thus there is no adjustment in H_r. Using equations (4) and (6) the price index in region d can be expressed as:

$$P_d = \frac{\sigma}{\sigma - 1} \left[\sum_{d=1}^{R} H_d T_{od}^{1-\sigma}\right]^{\frac{1}{1-\sigma}}$$

with $\sigma > 1$. Alternatively for region o:

$$P_o = \frac{\sigma}{\sigma - 1} \left[\sum_{o=1}^{R} H_o T_{od}^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (10)$$

For a given distribution of human capital across regions, we can derive from (5), (9), and (10) the equilibrium value for nominal wage, W_r:

$$W_o = \frac{\alpha}{\sigma} \sum_{d=1}^{R} \left[\frac{(H_d + L_d)\, T_{od}^{1-\sigma}}{\sum_o \left(H_o T_{od}^{1-\sigma} \right)} \right] \quad (11)$$

where $r \in \{1, .., d, .., o, .., r, .., R\}$ and R is the total number of regions.

In a long run, workers are mobile between regions. They migrate towards regions with the highest indirect utility. By migrating between regions workers equalise price real wages and, hence, indirect utilities among regions. Thus, long-run equilibrium is achieved when there are no differences in the indirect utility between regions and, hence, no incentive to migrate.

From (1), utility maximisation yields the following indirect utility function:

$$V_r = -\alpha \ln(P_r) + Y_r + \alpha \ln(\alpha - 1) \quad (12)$$

where Y is household income. Hence, one can derive the utility differential:

$$\triangle V_{od} = V_o - V_d = \alpha \ln(\frac{P_d}{P_o}) + (W_o - W_d) \quad (13)$$

Using equations (10) - (12), it is straightforward, that $\triangle V_{od}$ only depends on the share of mobile workers in region d and parameters of the model. As will be shown in Section 4.5, migration rate of mobile workers in turn depends on $\triangle V_{od}$. Thus, combining equations (13) and (19) creates a potential for a self-perpetuating agglomeration process, which in the literature is known as 'circular causality'. The reason is that immigrant inflows are likely to increase real wages in the destination region and outflows are likely to decrease real wages in the origin country. Because of this reverse causality, notice that it is likely to bias reduced-form models' estimates upwards.

Equations (10), (11) and (13) build core of a long-run equilibrium solution. Having data for exogenous variables and parameters values, we may solve the model and obtain a long-run equilibrium solution. The variables we are interested in are P_r, W_r, V_r and $\hat{H}_r$[7]. Hence, there are $4 \times R$ unknowns and $3 \times R$ equations. Generally, more variables than equations imply insolvability of the system. According to our definition of a long-run equilibrium, in equilibrium the utility differences $\triangle V_{do}$ must be zero between any pair of regions (13). Thus, we impose $\triangle V_{do} = 0$, $\forall r$ for each origin-destination pair instead of solving for endogenous values of V_r. The benefits

[7] This is not an exclusive list of all endogenous variables. We focus on these variables because of their strategic relevance determining inter-regional migration.

from doing so are twofold: (1) guarantee that the achieved solution is indeed a stable equilibrium ($\triangle V_{do} = 0 \Rightarrow$ no incentives to migrate); and (2) solvability of the model is achieved (number of endogenous variables = number of equations). Although, the model can be solved analytically for mobile workers share in each of the R regions, $\hat{H}_r$, the equilibrium expressions are rather involved and, therefore, are not presented here[8].

4.5 Inter-regional migration

On the basis of equation (13) we are able to calculate net migration flows for each region. If gross flows instead of net migration flows are required, one needs to rely on a different methodology. Choice of the exact methodology depends critically on assumptions regarding how individuals make migration decisions. Migration may be modelled as a sequential decision-making process. For example, an individual may first make the decision whether to move, followed by the choice of destination (for those who do move). This assumption underlies models considering the decision to move, separately from destination choice (for example *Levine* and *Zimmerman*, 1999), as well as models focusing only on destination choice of movers (*Bartel*, 1989). In the context of jointly modelling the decision to move and destination choice, an assumption of sequential decision-making leads to a nested logit framework, employed by *Frey et al* (1996). Alternatively, *Davies, Greenwood,* and *Li* (2001) argue that the decision to move and choice of destination cannot be separated. Their logic leads to a conditional logit model. In practice, some individuals likely jointly make the decision to move and destination choice, while others follow a more sequential process.

Modelling gross migration flows starts with an assumption that each mobile worker i from region o and has a location choice among R regions (including or excluding o). The migration function representing this migration process can be derived from utility maximising choices by individuals. Let a worker i in region o derive a utility $U_{i,od}$ from migrating to region d. Further we assume that the utility is a linear function of the relevant economic conditions pertaining to the origin (o) and destination (d) regions, denoted by a vector $V_{i,od}$. His migration choice results from a comparison of the perceived quality of life in various locations, which is formalised as a maximisation of the following objective function:

$$U_{i,od} = \beta_i V_{i,od} + \epsilon_{i,od} \qquad (14)$$

where β_i is a conforming vector of utility coefficients, which may vary depending on the identity of the worker[9], and $\epsilon_{i,d}$ is a stochastic component (random disturbance) capturing i's personal perception of the characteristics of region d. Assuming that there are R alternative destination regions, migration choices are determined from

[8] Detailed results are available upon request from the author.

[9] Traditionally in reduced-form models, the parameter β_i is kept constant across choices.

a comparison of $U_{i,od}$ across regions. If a worker chooses region d, then the utility $U_{i,od}$ is the highest among all R choices (i.e. $U_{i,od} > U_{i,or}$ for all $r \neq d$). According to *Davies et al.* (2001), when the choice d is made, the statistical model for the probability of moving from region o to region d can be represented as

$$\Pr\left(\Omega_r = d\right) = \Pr\left(U_{i,od} > U_{i,or}\right), \qquad \forall r \neq d \tag{15}$$

where Ω_r is a random variable that indicates the choice made. The model is made operational by a particular choice of distribution for the disturbances. Based on *McFadden* (1973), if and only if the R disturbances $\epsilon_{i,od}$ are independent and identically distributed with the Weibull distribution, then the probability of an individual at region o choosing region d (where $r = d$ for non-movers) is a conditional logit function:

$$\Pr\left(\Omega_r = d\right) = \frac{e^{\beta_i V_{i,od}}}{\sum_{r=1}^{R} e^{\beta_i V_{i,or}}} \tag{16}$$

It is convenient to rewrite equation (16) in terms of the individuals, which are migrating to region d as opposed to staying in region o:

$$\frac{\Pr\left(\Omega_r = d\right)}{\Pr\left(\Omega_r = o\right)} = e^{\beta_i\left(V_{i,od} - V_{i,oo}\right)} \equiv e^{\beta_i \triangle V_{i,do}} \tag{17}$$

Equation (17) states that the likelihood that an individual in region o migrates to region d rather than staying in region o is a function of the differentiated expected utility between migrating and staying. Equation (17) predicts the empirical migration rate $\frac{n_{od}}{n_{oo}}$, the number of people in region o who migrate to region d during a period, n_{od}, over the number of people in region o who stay, n_{oo}. Assuming that all individuals have the same utility function, we obtain the following empirical migration equation:

$$\frac{n_{od}}{n_{oo}} = \frac{\Pr\left(\Omega_r = d\right)}{\Pr\left(\Omega_r = o\right)} + \epsilon_{od} = e^{\beta_i \triangle V_{i,do}} + \epsilon_{or} \tag{18}$$

Taking a logarithmic transformation of equation (18), we obtain the following equation of a migration rate M_{or}:

$$\ln M_{od} = \ln\left[\frac{n_{od}}{n_{oo}}\right] = \beta_i \triangle V_{i,do} \tag{19}$$

On the basis of equation (19), we are able to calculate bilateral place-to-place migration flows, M_{od}, for each pair of regions. Using equations (10), (11), (13) and (19), the number of people moving from origin region o to destination region d can be expressed as a function of manufacturing price index, P_r, the wage rate in each region, W_r, and parameters of the model (α, σ, T_r).

5 Empirical implementation

This Section explains what kind of data and which parameters are required to run the NEG model, as well as refers to the main sources of the data used. This Section also explains how model's parameters have been estimated, and what kind of data is required to estimate the parameters. It is worth to mention that the approach we use in estimating the model's parameters is fundamentally different from the mainstream technique used in CGE (Computable General Equilibrium) models. Usually in CGE models parameters are calibrated. Although easily to apply, calibration adds lots of arbitrariness to the results obtained. The way we estimate model's parameters drawing on statistical data from the past, allows us to obtain both data-consistent and NEG model-conform estimates of the parameters.

5.1 Data

Data requirements of the NEG model and of the econometric estimations are fundamentally different. Therefore, two separate data sets are required - one for estimating coefficients of equation (22) and one for predicting migration flows using the NEG model. The estimation of equation (22) requires panel data of migration rate, M_{od}, and of all explanatory variables. The NEG model is less data demanding in quantitative terms, but is rather sensitive to the data consistency and accuracy. It requires only one observation (base year) data.

Currently such a detailed and consistent data set is is available only for three CEE accession countries - Estonia, Latvia and Lithuania[10]. Choice of the three North Eastern European accession countries has further advantages. First, due to the fact they were part of the former Soviet Union, Baltic states have a common migration history. The language and cultural differences between the three countries are rather low (Section 2.1). Second, because the three Baltic states are the most liberal ones among the CEE accession countries in terms of labour market access (see Section 2.2), we expect that after EU integration main migration flows will arise in these three countries.

The estimation of equation (22) requires panel data of migration rate M_{od}, and of all explanatory variables (H, T and X). *Regio* database (*Eurostat* 2004) provides annual bilateral migration data at NUTS II level. While invaluable, this information has several shortcomings. First, the time period covered in *Regio* is only 14 years (1989 - 2003). Even for this short time period, the *Regio* data is not complete - missing entries account for more than 25% of total observations[11]. To fill gaps in the data and to make the data internally consistent, we are reliant on national statistics'

[10] More CEE economies will be introduced in the future, when the necessary data becomes publicly available.

[11] Although a measurement error in the dependent variable does not lead to biased estimation results unless the measurement error is correlated with the exogenous variables, we decided to exclude years when the observed migration showed unrealistic spikes.

data, which we use extensively. Second, migration data is not internally consistent, e.g. for some years the number of emigrants leaving Latvia does not correspond to the number of immigrants arriving in Lithuania. For these two reasons, we do not pool all the data together but perform separate regressions for each country. Thus, the information contained in *Regio* together with the national statistics' data allows us to obtain migration data for 14 years building a panel of $14 * 2 = 28$ observations for each country.

In addition to migration data, national statistics provide data on employment, price index and industry output at NUTS II level. The two main shortcomings national statistics data has are: (1) discrepancies in countries' definitions used for the same variables (particularly when calculating manufacturing output); and (2) sectoral definition in the national statistics differs substantially among the three countries. Recognising these drawbacks in the data quality, we are able to extract internally consistent information only for 14 years (1989 - 2003) at NUTS II level. Thus, the information contained in *Regio* together with national statistics allows us to obtain a more or less consistent statistical data for H_r, L_r, and X_r for each NUTS II region building equally sized time-series with 14 observations in each.

The NEG model is considerably less data demanding in quantitative terms and requires a cross-section of NUTS II regions covering one year only. Because only one observation (base year) is used in the analysis, quality and consistency of the data is particularly important for the NEG model. In particular, the NEG model requires base year values for H_r, L_r and all model's parameters. Rest of the variables, such as migration rate, are calculated endogenously within the model. For this purpose we use data already prepared for the estimation of equation (22). Parameter values are obtained from separate regressions, which are explained in more detail in the following two Subsections (Section 5.3 and Section 5.4). In order to estimate transport cost parameters ϕ_{od} and $T_{od}^{1-\sigma}$, bilateral trade data is required. Eurostat External Trade Statistics (COMEXT CD-ROM) provides bilateral trade flows in SITC and NACE classification. It has an excellent time, country and sector coverage and is already available to the all CEE accession countries including the three Baltic states.

There are several data issues to address before performing estimations. A major difficulty arises from the definition of the traditional sector. According to the theoretical framework presented in Section 4, the major difference between sector X ('manufactured goods') and A ('traditional good') lies in the market structure and in the presence of scale economies: the 'traditional' sector should capture all homogeneous productions with constant returns to scale, while all tradable and differentiated productions with increasing returns to scale should be considered as 'manufactured goods'. Unfortunately, we do not have detailed sectoral data allowing such a classification. The simplest solution, therefore, is to consider agriculture as a proxy for 'traditional' production, so that the X sector captures all manufacturing activities. Another difficulty arises from the definition of the immobile factor. According to the theoretical framework, the major difference between the two factors L ('immobile factor') and H (labour) consists in the inter-regional mobility - while regional supply of the former is fixed, labour is mobile between regions. A further difference among the two factors is given by the sectoral employment - traditional sector uses only the immobile factor (L) in the production process, while manufacturing uses labour (H) as a fixed cost and the immobile factor for the variable cost. Unfortunately, we do not have detailed data allowing such a classification. The simplest solution, therefore, is to consider unskilled labour as a proxy for the 'immobile' factor, so that the H sector stands for skilled labour ('human capital').

5.2 Transport costs

One of the key parameters in the NEG model is the term $\phi_{od} = T_{od}^{1-\sigma}$, that *Baldwin et al.* (2003) cunningly refers to as the "phi-ness" of trade. We follow the standard practice in economic geography literature (*Head & Mayer* 2004) and assume symmetric bilateral trade barriers $\phi_{od} = \phi_{do}$. This assumption leads to a very simple estimator for ϕ_{od}:

$$\hat{\phi}_{od} = \sqrt{\frac{m_{od} m_{do}}{m_{oo} m_{dd}}} \tag{20}$$

where m_{od} is import of goods and services from an origin region o to a destination region d. The numerator in equation (20) requires only trade flow data expressed according to our industry classifications. The denominator factors are each region's "imports from self" (or, equivalently, "exports to self"). They are calculated as the value of all shipments of the industry minus the sum of shipments to all other regions (exports) (*Head & Mayer* 2004).

We estimate $\hat{\phi}_{od}$ for each pair of Baltic countries, where $\hat{\phi}_{od}$ is an inverse measure of transport costs, which ranges from zero to one. Table 4 reports the estimation results.

Table 4. Road distances and transport cost

Regions	D_{od} (*km*)	$\hat{\phi}_{od}$
R_{12}	297.1	0.316
R_{13}	608.6	0.187
R_{23}	290.8	0.349

D_{od} - road distances between capitals (*Route 66*).
Source: Own calculations based on *Comext* (2004).

Table 4 reports values of $\hat{\phi}_{od}$, which are obtained from estimating equation (20). Recalling that $0 < \phi_{od} < 1$ with 0 denoting prohibitive trade costs, the overall level of trade freeness appears to be rather low, even though we have calculated $\hat{\phi}_{od}$ for pairs of countries known for their high levels of formal trade integration, e.g. BAFTA - Baltic Free Trade Agreement.

Table 4 also reveals that trade costs between Estonia and Lithuania ($\hat{\phi}_{13}$) are considerably higher than trade costs among any other two regions[12]. These remarkable differences in estimated trade costs can be explained by regions' geographical location - the three regions are situated 'along a line' rather than 'in a circle'. Moreover, in our three region world (Estonia, Latvia, and Lithuania), transportation of goods between Estonia and Lithuania has not only the biggest average transport distances, in addition these goods always have to cross at least two borders[13]. Border-crossing costs are part of the general transport costs $\hat{\phi}_{od}$ and were comparatively high before integration. This might explain the remarkable differences in the estimated $\hat{\phi}_{od}$s.

Assuming these estimates are correct, the results presented in Table 4 should have implications on the scenario design, when turning to numerical simulations. It seems reasonable to assume that integration with the EU will reduce trade costs among the two peripheral border regions faster than among the core region and the peripheral regions. Practically this means that T_{13} should decrease faster than T_{od} estimates for any other pair of regions.

5.3 Econometric specification

As already explained, we want to estimate parameters in a specification, which is entirely derived from the theoretical NEG model. Thus, we need to derive an econometrically estimable equation, which contains the relevant parameters of the NEG model. The theoretical framework presented in Section 4 suggests an empirical specification characterised by the emigration rate as the dependent variable and, among the explanatory variables, wages and price indices in the origin and destination regions. In order to derive an estimable migration equation, we use the definition of $\triangle V_{do}$ and substitute equation (13) into (19). We obtain the following migration equation, where ϵ_{od} is a random prediction error:

$$\ln M_{od} = \beta \ln P_o - \beta \ln P_d + \beta W_d - \beta W_o + \epsilon_{od} \tag{21}$$

According to (21) migration rate M_{od} is increasing in origin region o's manufacturing price index, decreasing in destination regions d's price index, increasing in destination regions d's wage rate and decreasing in origin region o's wage rate. Equation (21) is a reduced-form model of migration rate, because we do not observe all explanatory variables. In particular a NEG model-consistent manufacturing price index, P_r, is not available in the data[14]. Moreover, (21) does not contain all parameters of the NEG model required to run the model. In fact, it contains only the consumer

[12] ϕ_{od} and $T_{od}^{1-\sigma}$ have been re-estimated for different years always obtaining comparable estimates to those ones presented in Table 4.

[13] Bilateral trade between any other pair of regions involves crossing only one border.

[14] Consumer price index (CPI) is readily available in statistical data. However, because of the way the CPI is constructed is completely different from the manufacturing price index in the NEG model (10), it is not consistent and cannot be used in the estimations.

preference parameter α. Thus, (21) requires further transformations before it can be estimated structurally. First, because manufacturing price index P_r is not observable in the data, it needs be substituted out. We use (10) to obtain a migration equation without the manufacturing price index. Next we introduce two new variables, Λ_d and Λ_o ($\Lambda_d = \sum_{d=1}^{R} H_d \hat{\phi}_{od}$ and $\Lambda_o = \sum_{o=1}^{R} H_o \hat{\phi}_{od}$). The new variable, Λ_d, contains a transport cost component, ϕ_{od}. Although, at this stage we still do not know the numerical values of T_{od}[15], this is not an obstacle for estimation of (22), because we can straightforwardly plug in the estimated values of $\hat{\phi}_{od}$ without an exact knowledge of the inter-regional trade cost parameter T_{od}. Second, we use (9) and substitute out W_r ($W_r = \frac{1}{\sigma-1} X_r$, $\sigma \neq 1$). This substitution allows us to obtain a coefficient containing the elasticity parameter σ_r. Without σ_r we would not be able to separate α_r from σ_r in the coefficients β_2 and β_3. And, more importantly, the manufacturing output data is more reliable than wage data in the CEE transition economies, where a considerable part of wages are still being paid in 'envelopes'. Last, we impose restrictions on the structural parameters implied by the theoretical NEG model: $\sigma > 1$ and $\alpha > 0$. These intervals correspond to the following coefficient restrictions: $\beta_2 < 0$, $\beta_3 < 0$ and $\beta_4 > 0$, $\beta_5 > 0$. Because β_2 and β_3 are strictly negative, migration rate is increasing in Λ_d and decreasing in Λ_o. Thus, we obtain the following linear estimable migration equation:

$$\ln M_{od,t} = \beta_1 + \beta_2 \ln \Lambda_{o,t-1} - \beta_3 \ln \Lambda_{d,t-1} + \beta_4 X_{d,t-1} - \beta_5 X_{o,t-1} + \beta_6 \epsilon_{or,t} \qquad (22)$$

where $\beta_2 = \frac{\alpha_d}{1-\sigma_d}$, $\beta_3 = \frac{\alpha_o}{1-\sigma_o}$, $\beta_4 = \frac{1}{\sigma_o-1}$, $\beta_5 = \frac{1}{\sigma_d-1}$ are coefficients to estimate, and $\Lambda_o = \sum_{o=1}^{R} H_o \hat{\phi}_{od}$, $\Lambda_d = \sum_{d=1}^{R} H_d \hat{\phi}_{od}$, X_o and X_d are the explanatory variables. Note that all explanatory variables in (22) are consistent and available in the data. Thus, having time series data for M_{od}, H_r and X_r, and estimates for $\hat{\phi}_{od}$ we are able to estimate (22). This structural specification provides consistent estimates of key parameters of the NEG framework (α_r and σ_r). α_r is a parameter, which determines consumers' preferences for manufactured goods and σ_r is the elasticity of substitution.

Equation (22) states that the migration rate M_{od} from an origin region o to a destination region d is a function of origin and destination regions' share in total supply with mobile workers, H_r, a function of origin and destination regions' manufacturing output, X_r, and a function of inter-regional transport costs, ϕ_{od}. Clearly, unobservable economic and non-economic characteristics of regions such as amenities also play an important role in a migration decision. To capture these effects we include a set of dummy variables for potential destinations (region fixed effects). We will assume throughout the estimations that error terms are not correlated with explanatory variables and fixed effects, and are not serially correlated, so the fixed-effects estimation is not biased. We follow the standard practice in migration models with cross-sectional data (*Boeri & Brücker* 2001) and pool all origin regions together. Thus, we obtain

[15] In order to calculate values of $T_{od}^{1-\sigma}$ we would require estimates of σ_r.

a panel[16] covering (14 years) $\times$ ($R-1$ regions). The two-dimensional nature of the panel data allows us to exploit both the variation between regions and time periods in the data for the estimation of the parameters of the migration function.

Different estimators applied in the literature use different sources of variation and, hence, yield different results. Most often used estimators could be grouped as follows: (i) *Traditional estimators*: pooled OLS, *within*, SUR, random effects (GLS) estimators, and the iterated GLS (Maximum Likelihood) estimator; (ii) *Instrumental variable estimators*; and (iii) *GMM estimators* (*Arellano* & *Bover* 1995). The pooled OLS estimator employs both sources of variation (time and cross-sectional) in the data, albeit not efficiently. The main feature of the pooled OLS estimator lies in its uniform treatment of all cross-sections. In other words, it neglects the individual heterogeneity amongst the different countries such as, geography that may have an important impact on migration. On the other hand, the *within* and SUR estimators allow for differences among cross-sections captured by the individual specific intercept term. In doing so, these estimators focus only on the within variation in the data as the name of the former estimator suggests. The main difference between these two estimators lies in different restrictions imposed on the covariance matrix of the error term. The *within* estimator imposes the same covariance matrix for all cross-sections, whereas the SUR estimator relaxes this assumption by allowing for different covariance matrices across the cross-sections with possible correlation between the individual specific error terms. When the imposed assumptions of the GLS estimators are fulfilled, the GLS estimators are superior in terms of efficiency when compared to the other traditional types of estimators. This is achieved by the optimal weighting attached to the within and between variations in the data.

In dynamic panel models, which model sluggish adjustment by considering lags of the dependent variable, the traditional estimators are subject to a simultaneous equation bias. This bias is caused by the presence of the lagged dependent variables among the explanatory variables. This simultaneity bias can arise for the following reasons: (a) presence of the individual specific effects, which will definitely be correlated with the lagged dependent variable; (b) autocorrelation in the error term of the migration function that results in correlation between the lagged dependent variable and the regression disturbances. The GMM estimator developed by *Arellano* & *Bover* (1995) addresses this issue by employing both the first differences as well as the levels equations by specifying the appropriate sets of instruments for either types of equation.

In summary, given the different properties of the estimators available, we would expect that for panels with rather small time series dimensions and rather large cross-sectional dimensions, the GMM estimator of *Arellano* & *Bover* (1995) will be superior to the rest of the estimators discussed above. In particular, we would expect the traditional estimators to perform rather poorly due to the unresolved simultaneous

[16] Various authors refer to longitudinal data, cross sectional data, panel data, and cross-sectional time-series.

equation bias. When the opposite is the case, i.e. relatively large time series and rather small cross-sectional dimensions of the panel, the motivation for using the GMM estimators is less obvious as these were primarily designed for the former case. In addition, the inflation of the moment conditions may somewhat worsen their performance in panels typically considered in macroeconomic studies. On the other hand, either the *within* or SUR estimator is expected to have a comparative advantage over the rest of the estimators as the simultaneity bias is likely to be smaller in the magnitude than in panels with a short time dimension.

In order to decide which estimator to use, we test first for individual heterogeneity using the *Breusch-Pagan* test, which compares pooled OLS to the random effects model, and reject the null of no heterogeneity. This leads to the question whether unobserved region-specific effects are correlated with the explanatory variables. We use Hausman's specification test and reject the null of zero correlation, thus a fixed effects model is preferred to a random effects model. As expected, given the relatively small cross-sectional dimension of our panel ($R = 3$) compared to the time dimension ($t = 14$), the specification tests favour neither a pooled OLS, nor a random effects model. Therefore, we chose the *fixed effects* estimator with dummy variables (which is equivalent to modelling the *within* estimator).

Given the chosen specifications, an empirical complication - the problem of endogeneity arises. If one might assume that migration flows are relatively too small to induce reverse causality (the impact of immigrants on wages is the focus of analysis in Section 6), however there might still exist confounding factors that influence contemporaneously manufacturing output, wages and migration flows. For instance, negative demand shock may drive wages down and, at the same time, increase emigration. To (partly) get around this problem we use instrumental variables estimation with lagged values of manufacturing output and stock with mobile workers as instruments. Thus, we assume that migration choices at date t are determined from a comparison of V_r across regions at date $t-1$. We restricted the number of lags to one in order not to lose further time-series observations. We have to assume that instruments are predetermined, and immigrant inflows and confounding factors in residuals only affect contemporaneous and future production, and migrant stocks. We also have to assume that residuals are not autocorrelated, otherwise the instruments will be invalid.

5.4 Estimation results

Equation (22) has been estimated for the three selected CEE accession countries: Estonia, Latvia and Lithuania using *random effects* and *fixed effects* estimators. Table 3 reports the estimation results. The first two columns in Table 3 report estimates from *random effects* specification. Columns 3 and 4 report estimates from *fixed effects* specification. Most of the parameters presented in Table 3, converge toward consistent values. Although, the estimation results presented in Table 3 show substantial differences among the three Baltic countries, the estimated coefficients always have the same order of magnitude. All estimated coefficients have the expected signs. When region-specific effects are used, coefficients' signs do not change.

According to $t-$statistics all explanatory variables are significant. However, the impact of different explanatory variables on migration rate is rather differentiated. First, migrants respond stronger to the manufacturing output than to the log of region's share in the total workforce. Second, the impact of the two presented explanatory variables ($\log \Lambda_d$ and X_d) have opposite signs. Destination region's manufacturing output, X_d, is positively related to the migration rate - the higher is the manufacturing output in the destination region, the higher utility utility and higher incentives to migrate to the particular region. The same is true for destination region's share in human capital weighted by inter-regional transport costs, $\log \Lambda_d$. The higher is region d's share in total labour supply, the higher is the migration rate to the particular region. These coefficients are positive and are in line with our expectations (columns $\log \Lambda_{d,t-1}$ in Table 3).

Table 3. Dependent variable: log migration rate

	Random effects		Fixed effects		
	$\log \Lambda_{d,t-1}$	$X_{d,t-1}$	$\log \Lambda_{d,t-1}$	$X_{d,t-1}$	R^2
R_1	-0.58	1.43	-0.82	1.14	0.61/0.64
	$(3.17)^{\dagger\dagger}$	$(1.95)^{\dagger}$	$(4.75)^{\dagger\dagger}$	(0.85)	
R_2	-0.78	2.56	-0.84	3.18	0.67/0.79
	$(9.85)^{\dagger\dagger}$	$(3.53)^{\dagger\dagger}$	$(8.97)^{\dagger\dagger}$	$(3.59)^{\dagger\dagger}$	
R_3	-0.61	2.31	-0.69	2.58	0.73/0.78
	$(3.56)^{\dagger\dagger}$	$(5.08)^{\dagger\dagger}$	$(3.81)^{\dagger\dagger}$	$(5.37)^{\dagger\dagger}$	

Panel data estimates, with heteroscedasticity corrected t-values in parentheses. †significant at 5%, ††significant at 1%. All variables lagged by one year. $R_1 = Estonia$, $R_2 = Latvia$, $R_3 = Lithuania$.

From the estimated coefficients we may yield consistent estimates of the NEG model's parameters. From each estimation we can calculate two parameters, α_d and σ_d. In order to obtain all the necessary parameters for each region, we need to repeat the estimation of (22) R times (in our case three). Because of econometric considerations discussed in Section 5.3 we use the *fixed effects* estimates presented in columns 3 and 4 to calculate the parameters values.

Regarding magnitudes of the numerical estimates, our results are not straightforwardly comparable with existing migration studies, because of obvious differences in the explanatory variables. Our migration model is specified with a focus to estimate parameters of the NEG model, while in the literature it is usually estimated for prediction purposes. The overall statement in the literature is that a common labour market will not initiate massive labour migration, but peak levels of migration may be plausible during the first years. Accordingly, up to 1.2-2.5 % of the CEE-10's current population is expected to migrate to EU-15 countries in the medium and long run (10-30 years) (*Bauer* & *Zimmermann* 1999, *Fertig* 2000, *Boeri* & *Brücker* 2001, *Kielyte* & *Kancs* 2002). In the case of the 2.5 % estimate proving accurate the actual number of migrants would correspond to the present population of Ireland, or somewhat more than the Swedish population. For the comparison purposes we express the estimated coefficients in terms of elasticities of migration. Elasticities of migration are of particular interest, because they form backbone in most of the migration studies. The migration elasticities we have obtained are somewhat lower than those estimated in the literature. Deviations among previous studies and our estimates might be caused by differences among estimators used and by statistical problems, small number of observations being one of them.

6 Predicting migration flows

This Section is devoted to simulating migration flows in selected CEE accession countries - in Estonia, Latvia and Lithuania in the years following the CEE's integration with the EU. Using the NEG model developed in Section 4, the estimated parameters from Section 5, and the *Regio* data from 2003, we are able to calculate an equilibrium solution for different levels of trade costs, which should mirror different levels of EU integration.

In the presented simulation experiments we show that migration might induce effects in the system, which work in a reverse direction (from the left-hand-side towards the right-hand side variables). Thus, migration not only absorbes distortions caused by a short-run disequilibrium in the economy, it also induces changes in the system. The standard reduced-type approach, because of its partial equilibrium character, cannot capture these adjustments. An estimation without considering adjustments in the explanatory variables, induced by workers migration and industry relocation, results in biased results. Moreover, because of a reverse causality, reduced-form models' estimates will likely be biased. The NEG model applied in this Section captures both, the migration induced changes in the explanatory variables, and the wage and utility induced adjustments in migration.

6.1 Base run

We use the *Regio* data from 2003 for each of the three Baltic states to implement and run the NEG model. The model is implemented by the General Algebraic Modelling System (GAMS) and solved using the CONOPT and MINOS solvers (*Brooke et al* 1988). The (dis)equilibrium solution we obtain is used as a benchmark in all simulation exercises throughout rest of the paper. In a second step (Section 6.2) we then compare the obtained estimates with the 'base run' and calculate differences in the regions' stock with mobile workers.

One feature sets the model employed here apart from the traditional specification of reduced-form migration functions. The great majority of the empirical studies explains migration flows by income and employment variables and (lagged) migration stocks which should capture network effects. This specification implicitly assumes that a long-run equilibrium relationship between migration flows and the explanatory variables exists. In terms of time-series econometrics this requires that all variables are either stationary or integrated of the same order. This is not very plausible, since the migration flow variable is (largely) the first difference of the migration stock variable, such that it is unlikely that both the migration flows and the migration stocks are I (1) variables. In contrast, the NEG model employed here assumes that an equilibrium relationship between migration stocks and macro-economic variables such as utility differences exist in the long run[17].

We start with solving the model for a long run equilibrium, which allows us to obtain values of all endogenous variables, such as prices, manufacturing output, wages, sectoral employment and bilateral migration flows (if any) for each region. We use the model developed in Section 4 and fit the *Regio* data and the estimated parameters' values. Running the model we obtain values of all endogenous variables, four of which are reported in Table 5. For convenience, this solution is called the 'base run'.

Table 5. Base run estimates, P_r, W_r, V_r

Region	P_r	W_r	V_r
R_1	1.6238	0.1096	0.0001
R_2	1.4747	0.1428	0.0659
R_3	1.4265	0.1539	0.0884

Source: NEG model simulations based on *Regio* data. For assumptions see Section 4.
$R_1 = Estonia$, $R_2 = Latvia$, $R_3 = Lithuania$

According to figures presented in Table 5 the highest indirect utility is in region R_3 (Lithuania), and the lowest in region R_1 (Estonia). These differences are easily explained by the other two variables (P_r, W_r) presented in Table 5. Estonia has the highest manufacturing price index and the lowest wage rate. Lithuania, on the opposite, faces the lowest manufacturing price index and the highest nominal wage rate. Differences in the manufacturing price index and the wage rate yield differences in the indirect utility. Thus, our base run calculations are in line with the NEG theory *Krugman* (1991).

As next we want to consider how far our base run estimations are from the observed data. Table 5 reports that indirect utilities are not equal in all three regions

[17] The hypothesis, whether migration stocks or flows and the explanatory variables are characterised by an equilibrium relationship, or, in more technical terms, are cointegrated, can be tested statistically (see e.g. *Chiswick* 2005).

in the 'base run' and the utility differences ($\triangle V_{od}$) are different from zero for all pairs of regions. Hence, according to our model, in the base year we do not observe a long run equilibrium between the stock with mobile workers and the explanatory variables. The transition to the long-run equilibrium occurs through workers' migration among regions. According to the underlying economic geography theory *Krugman* (1991), an additional migrant may work as an agglomeration force (when further workers follow), or as an equilibrium force (when the incentive to migrate declines).

Having estimates of P_r, W_r and V_r it is straightforward to calculate how many mobile workers should move among regions to achieve the long run equilibrium between the stock with mobile workers and the 'explanatory variables'. Subtracting short-run equilibrium $\hat{H}_r$ values from the long run equilibrium $\hat{H}_r$ values we obtain a net migration rate for each region. This migration rate, $\hat{M}_r^{BR}$, gives how many mobile workers would move from low wage and high price regions to high wage and low price regions in order to establish a long run equilibrium. The obtained results are presented in Table 6.

Table 6. Predicted and observed net migration

Regions	$\hat{M}_r^{BR}$	$\triangle \hat{H}_r^{BR}$	M_r^{2003}	$\triangle H_r^{2003}$
R_1	-1.090	-1492.6	0.381	521.2
R_2	-0.627	-1488.9	-0.799	-1898.4
R_3	0.859	2981.5	0.397	1377.2

Source: NEG model simulations based on *Regio* data.
$BR-$predicted, $2003-$observed, M_r-migration rate.
$R_1 = Estonia$, $R_2 = Latvia$, $R_3 = Lithuania$

For comparison purposes, Table 6 also reports the observed migration, which is taken from the data (see Section 5.1 for explanation). Migration rate, M_r^{2003}, is expressed per 1000 workers and $\triangle H_r^{2003}$ reports changes in the stock with mobile workers. Comparing our base run estimates ($\hat{M}_r^{BR}$, $\triangle \hat{H}_r^{BR}$) with the observed values (M_r^{2003}, $\triangle H_r^{2003}$) provides an intuition about the magnitude of the prediction error. In the case of Latvia and Lithuania our estimates are pretty close to the true values (rows R_2 and R_3 in Table 6). NEG model's estimates are less precise in the case of Estonia. Whereas according to our estimates a long-run equilibrium relationship in the regional labour markets will be established if 1493 workers from Estonia would migrate to Lithuania, the data from 2003 reveals a reverse migration. These differences show that the numerical values should not be overemphasised, but have to be seen in a context of model's assumptions. According to the NEG model, the only way how to compensate differences in the indirect utility (column V_r in Table 5), is to move from low-wage to high-wage regions. In reality, because of language, cultural, climatic and many other differences, workers might prefer to stay in or even to move to low-wage regions.

6.2 Simulation exercises

We follow a modelling approach, which dates back to 1950-ties and model European integration and increased labour mobility as a reduction in transport costs. According to their approach, in order to set up an integration scenario of decreasing transport costs one requires: (1) magnitude of the real transport costs at a benchmark, and (2) estimates of changes in the transportation costs.

The real transport costs at benchmark have already been estimated in Section 5.2. Reliable estimates on transportation cost changes related to the European integration are not available in the literature yet. To overcome this deficiency, we construct several hypothetical scenarios, where transport costs between the peripheral border regions (T_{13})[18] are reduced in 10% steps up to 60%. Although, we have no particular reason to believe that European integration will induce transport costs reduction in this particular order of magnitude or in such an asymmetric way favouring the peripheral regions, these results should help us understand what type of labour market effects we should expect from further European integration.

In order to calculate new equilibrium values, we proceed in the following way. First, we exogenously change transport costs (T_{13}). Solving the model for a short-run equilibrium we obtain a solution with inter-regional differences in the price index, wages and indirect utility. Clearly, this is not a stable long-run equilibrium solution, because these differences contain an incentive for relocating. Since we are interested in a long-run equilibrium solution between the stock with mobile workers and the 'explanatory variables', in a second step we ask, what should be the new regional $\hat{H}_r$ in order to obtain the same level of indirect utility in each region? In other words, we fix $\triangle V_{do} = 0$ for all pairs of regions, and solve the model for equilibrium values of $\hat{H}_r$. Net migration rate is then calculated as $\hat{M}_r = H_r - \hat{H}_r$[19].

Table 7 reports simulation results, where $T = 100\%$, $T = 80\%$, $T = 60\%$, and $T = 40\%$ are the respective levels of transport costs. Columns 2-5 report the predicted migration flows as a percentage per 1000 mobile workers. Considering estimates in Table 7 it is straightforward to identify that different levels of transport costs lead to globally consistent[20] estimates of migration flows. While our results show substantial differences in the migration rate among the three Baltic countries, the total net migration flows (immigration minus emigration) sum up to zero in each period. Generally, the two peripheral regions (R_1 and R_3) are net winners in terms of mobile workers and industry, if the economic integration follows a pattern we assumed

[18] Transport costs between the peripheral border regions T_{13} are of particular interest, because they are much higher in the pre-accession state and because the two countries have a direct border with rest of the EU. We expect that integration with the EU will reduce transport costs between these two regions to a higher extent than between any other two regions.

[19] Analytically we are able to calculate a long-run equilibrium solution for regions' share with mobile factor. Empirically, in a R-region case with region-specific parameters it turns out to be impossible to solve the model for a long-run equilibrium in just one step.

[20] Zero net balance, when all regions weighted by their population and summed up.

in the simulations. When trade costs are reduced, Lithuania seems to be the biggest gainer in terms of labour force. It steadily receives mobile workers from other regions throughout the simulations (row R_3 in Table 7). Accordingly, the immigration rate in Lithuania, $\hat{M}_3$, is continuously increasing from 7.071% when transport costs with Estonia are reduced by 20% to 7.463% when T_{13} is reduced by 60%.

Latvia seems to be the largest looser from the modelled integration process. These results are not particularly surprising given that transport costs are reduced asymmetrically favouring the two peripheral regions. The emigration rate from Latvia, $\hat{M}_2$, is continuously increasing from 3.903% when transport costs are reduced by 20% to 11.576% when T_{13} is reduced by 60% (row R_2 in Table 7). Estonia is eventually the most interesting region from an analytical perspective and might be subject to predictions of the NEG theory. Our simulations show that the relationship between the level of regional integration in terms of trade costs and the regions' stock with mobile factors is non-linear and even non-monotonic. Initially, trade cost reduction among the two peripheral regions (region R_3 large and region R_1 small in relative terms) gives rise to agglomeration of workers and forms in the largest region (Lithuania). At the beginning of the simulated integration process the smallest region, R_1, looses more than 11% of its mobile workers. When trade costs fall below some critical point (sustain point) a diversified equilibrium becomes sustainable and Estonia - the smallest and the most peripheral region starts to increase its share with mobile workers (row R_1 in Table 7).

Table 7. Transport costs and net migration

Regions†	$\hat{M}_r^{20}(\%)$ $T_{13} = 80\%$	$\hat{M}_r^{40}(\%)$ $T_{13} = 60\%$	$\hat{M}_r^{60}(\%)$ $T_{13} = 40\%$
R_1	−11.018	−9.062	1.789
R_2	−3.903	−5.217	−11.576
R_3	7.071	7.354	7.463

Source: NEG model simulations based on *Regio* data.
† $R_1 = Estonia$, $R_2 = Latvia$ and $R_3 = Lithuania$
Positive values = immigration, negative = emigration.

The decreasing share with mobile workers at the beginning and increasing at some lower level of inter-regional trade costs is best explained considering the economic forces driving the model. Exogenous changes in transport costs are absorbed through changes in prices, wages, quantities produced, consumed and quantities transported, until a new equilibrium is reached. Because of these changes in the 'explanatory variables', the indirect utility is no longer equal among regions. Workers and firms become an incentive to move towards regions with a higher utility/profit. Firms' entry (workers' immigration) in turn actuates further movements in the system. We identify three such effects of firms' entry (workers' immigration)[21]. The first arises from the *price-index effect*: an extra firm lowers the industry price index which reduces the demand facing each existing firm (the demand and marginal revenue curves shift downwards). This competition effect, captured though the number of firms in equation (4), reduces profits encouraging the stability of a diversified equilibrium. The second effect is a *demand* or *backward linkage*. New workers raise demand for local varieties (demand and marginal revenue curves shift upwards), and this tends to raise profitability and so encourage more firms to follow. Additional firms in turn raise demand for labour in the destination region. This puts an incipient upward pressure on local wages, which encourages more workers to migrate. As we have seen, entry by a new firm lowers the price index in the destination region. But this induces a third effect, since it reduces the cost of living for workers and so tends to raise real wages in the destination region. The resulting migration restores inter-regional equality of real wages, which means that the nominal wage must fall. The result is a *cost* or *forward linkage* (which shifts the average and marginal cost curves downwards) and so raises profitability encouraging further firms to entry. Because, these two effects (*price-index effect* and the *demand* & *cost linkages*) work in opposite direction, a regional integration associated with a decrease in transaction costs between regions has an ambiguous effect. Numerical simulations, as presented in Table 7, offer a useful insight in the relationship among these forces.

6.3 Policy recommendations

From the simulations presented in Sections 6.1 and 6.2 we may conclude that the impact of economic integration in Central and Eastern Europe will depend crucially on the two effects, the *price-index effect* and the *demand* & *cost linkages*. Because, these two effects go in opposite direction, a regional integration that leads to a decrease in transaction costs between regions has an ambiguous effect on relocation of mobile labour force. Depending on the current state of the particular region, it may give rise to either immigration or emigration. Based on these results, it is impossible to derive an universally applicable regional migration policy.

The NEG model presented in Section 4 is solvable analytically in sence that we can find closed-form solutions for all endogenous variables, including the stock with mobile workers, $\hat{H}_r$, in each region. However, the analytical expressions of $\hat{H}_r$ are extremely involved and, hence, not particularly tractable. Therefore, the derivation of a formal equilibrium relation between size of regions & inter-regional transport costs and workers & firms' location is left beyond the scope of the paper. Instead, an insight into the equilibrium relation between size of regions & inter-regional transport costs and workers & firms' location is offered by performing numerical simulations. Closer considering results of numerical simulations presented in Section 6.2 we found that

[21]Note that the number of varieties produced equals the number of firms located in that region, which is linked one to one to the number of workers. Thus $H_r = N_r$.

economic integration that reduces transaction costs between regions will be successful in attracting firms and workers to a region only if the market size of the region is not too small and if existing transaction costs between the regions are not too low. Another way to say this is that an 'empty' place even if it is at the crossroads of large regions can not become an industrial base: a large enough local market is necessary. Also, to be attractive as a location that saves on transaction costs, those costs must be high enough. Eventually, Latvia is not successful in attracting firms and mobile workers because these costs are not high enough (they are several times lower than among Estonia and Lithuania).

In line with *Kancs* (2005), These numerical results imply that regional migration policy should be designed individually taking into account each region's individual characteristics (in terms of physical geography, market size and infrastructure) and policy objectives. The assistance in the development of declining regions should help the CEE accession countries in moving from low-employment equilibrium, high regional dependency (from transfers) to the higher rates of activity and more balanced regional labour markets.

7 Conclusions

In this paper we develop an analytically solvable and a structurally estimable economic geography model and apply it to predict migration flows for the period following the European integration. Model's parameters are estimated using a structural migration equation, which is derived entirely from the theoretical NEG model.

Although, the New Economic Geography has a history of almost two decades, its role in migration literature is still rather limited.[22] We use an estimated NEG model to demonstrate the tremendous potential of NEG's structural nature. We show that the general equilibrium nature of the NEG framework might be of high interest in studying migration questions, because it is able to cope with the problems caused by endogeneity and reverse causality between right- and left-hand side variables. The NEG model is able to cope with these bias by endogenising both left-hand side *and* explanatory variables.

The downside of the current approach is that a structural model *per se* does not guarantee a better fit - certain reduced-form specifications might still perform better in terms of forecasting performance. Therefore, we urge for more research, both methodological and empirical, be devoted to estimating and testing NEG models in predicting location of firms and workers. Future expectations also play a significant part in migration decisions. Expecting improvements in the home country may delay emigration decision or ultimately erase the idea. This issue has not been considered in the current paper and is a promising avenue for future research.

Our empirical findings advocate that there is enough evidence to predict a selective

[22] *Crozet* (2004) is one of the notable exemptions.

migration among the three Baltic states, when integration with the EU (modelled as a transport cost reduction) takes place. According to our results, labour mobility in the Baltic countries is sufficiently low to make the swift emergence of a core-periphery pattern very unlikely at NUTS II geographical level. As far as the economically motivated migration depends on differences in the level of prosperity between regions (indirect utility in our model), such differences will become less marked, as Europe becomes more integrated. Lithuania, having the largest internal market and being the most peripheral at the beginning of integration, turns out to be the biggest winner in terms of industry and mobile workers.

References

[1] Arellano, M., O. Bover (1995), Another look at the instrumental-variable estimation of error-components models, Review of Economic Studies, 58, pp. 277-97.

[2] Baldwin R., R. Forslid , Ph. Martin, G. Ottaviano and F. Robert-Nicoud. (2003), Economic geography and public policy. Princeton, N.J., Princeton University Press.

[3] Bartel, Ann (1989), Where Do the New U.S. Immigrants Live? Journal of Labor Economics, 7, 371-391.

[4] Bauer, Thomas, and Zimmermann, Klaus F., (1999), Occupational Mobility of Ethnic Migrants, Discussion Paper No. 58, IZA, Bonn, Germany, September.

[5] Boeri, T., H. Brücker et al. (2001), The Impact of Eastern Enlargement on Employment and Labour Markets in the EU Member States, European Commission, DG Employment and Social Affairs, Brussels.

[6] Brooke, A., David Kendrick and Alexander Meeraus (1988), GAMS: a User's Guide, New York: The Scientific Press.

[7] Chiswick, Barry R. (2005), The Economics of Immigration. Edward Elgar Publishing, Cheltenham, UK.

[8] Commission of the European Communities (2004), First Annual Report on Migration and Integration, European Commission, Brussels.

[9] Crozet, M. (2004), Do migrants follow market potentials?, Journal of Economic Geography. 4, no. 4, pp. 439-458.

[10] Davies, Paul, Michael Greenwood, and Haizheng Li. (2001), A Conditional Logit Approach to U.S. State-To-State Migration, Journal of Regional Science, 41, 337-360.

[11] DeNew, J., K.F. Zimmermann (1994), Native Wage Impacts of Foreign Labor: A Random Effects Panel Analysis, Journal of Population Economics, 7, pp. 177-192.

[12] Dustmann, C. (1996), Return Migration: The European Experience, Economic Policy, 22, pp.213-242.

[13] Eurostat (2004), Central European countries' employment and labour market review, Luxembourg.

[14] Faini R., de Melo J., Zimmermann KF (1999), Migration: The Controversies and the Evidence, London: Cambridge University Press for CEPR.

[15] Fertig M. (2000), The Economic Impact of EU-Enlargement: Assessing the Migration Potential, paper presented at the Workshop at University of Kent at Canterbury, 11-12 February, 2000.

[16] Fertig, M., C.M. Schmidt (2001), First- and Second-Generation Migrants in Germany – What Do We Know and What Do People Think, IZA Discussion Paper 286, IZA, Bonn.

[17] Frey, William, Kao-Lee Liaw, Yu Xie, and Marcia Carlson (1996), Interstate Migration of the U.S. Poverty Population: Immigration Pushes and Welfare Magnet Pulls, Population and Environment, 17, 491-536.

[18] Harris, J.R., M.P. Todaro (1970), Migration, Unemployment and Development, A Two-Sector Analysis, American Economic Review, 60, pp. 126-142.

[19] Hatton TJ (1995), A Model of UK Emigration, 1870-1913. Review of Economics and Statistics, 77:407-15.

[20] Head, Keith and Thierry Mayer (2004), The Empirics of Agglomeration and Trade, in Handbook of Regional and Urban Economics, vol. 4.

[21] International Organisation for Migration (IOM) (2005), World Migration 2005 Costs and Benefits of International Migration, International Organisation for Migration, Research and Publications Unit, Geneva.

[22] Kancs, D. (2005), Migration in the SEE and Integration with the EU: A Policy Perspective, Economics and Econometrics Research Institute - EERI, Brussels.

[23] Kielyte, J., Kancs, D. (2002), Migration in the Enlarged European Union: Empirical Evidence for Labour Mobility in the Baltic states, in Journal of Baltic Studies, Vol 33, No. 3, pp. 259-279.

[24] Krugman, P. (1991), 'Increasing Returns and Economic Geography', Journal of Political Economy, 99, 3, 483–99.

[25] Krugman, P., A. Venables (1995), Globalization and the Inequality of Nations, Quarterly Journal of Economics, 110, 857-880.

[26] Levine, Phillip and David Zimmerman (1999), An Empirical Analysis of the Welfare Magnet Debate Using the NLSY, Journal of Population Economics, 12, 391-409.

[27] McFadden, D. (1973), Conditional Logit Analysis of Qualitative Choice Behavior, in P. Zarembka, ed., Frontiers in Econometrics (Academic Press, New York).

[28] OECD (2003), Labour Market and Social Policies in the Baltic Countries, Paris.

[29] Forlsid, R., G. I. P. Ottaviano. (2003), An analytically solvable core-periphery model, Journal of Economic Geography 3(3):229-240.

[30] Sjaastad L.A. (1962), The Costs and Returns of Human Migration, Journal of Political Economy, 70:80-93.

[31] Straubhaar T., Zimmermann K.F. (1993), Towards a European Migration Policy, Population Research and Policy Review, 12:225-41.