

Economic Geography and intra-CEE migration

d'Artis Kancs

March 2005

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. 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.

1 Introduction

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. Although, it is an integral part of the *acquis communautaire*, in the light of recent European Union integration, labour migration issues have attracted huge public attention and remain highly controversial at all political and societal levels.

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

The high relevance of the topic has generated a large body of theoretical and empirical literature that attempts to predict the size of a possible labour migration in the years following the accession. Most of these studies are based on reduced-form econometric models and, in addition, focus solely on East-to-West migration. In this paper we argue that the reduced-form approach, especially when applied to economies in transition, where explanatory variables are due 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 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, in which part of Europe the recent EU integration might actuate significant 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 simulations results and policy recommendations are presented in Section 6. Section 7 concludes.

2 Migration in CEE

This Section focuses on the main past as well as more recent patterns of international migration concerning the new EU members and the accession countries. It also identifies migration policies, which might 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 analysis is followed by an identification of major directions of the future flows, as well as of main origin and destination countries.

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 CEE rather than East-to-West migration will be at the heart of matters after the EU enlargement. This can be 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 socialist regime.

New EU member states' migration history is remarkably different from that of the EU-15 member states. After the turmoil of post-war migration, which was the consequence of new post-Yalta World and 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 & Kanacs 2002*).

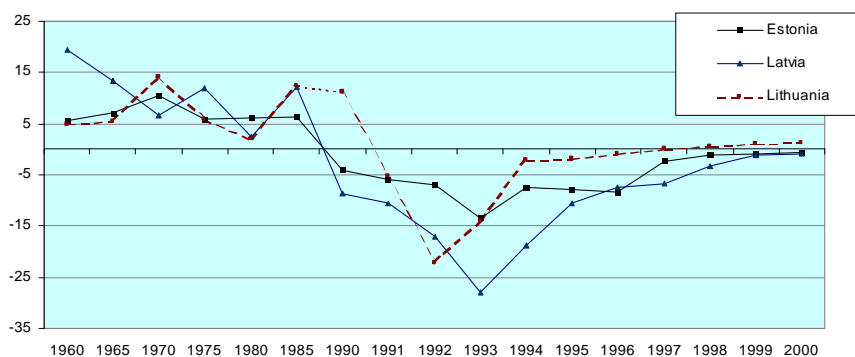


Figure 1: Net Migration in the Baltic States in Thousands, 1960-2000

In contrast, in Central and Eastern Europe the migration was deemed to be a political issue and, therefore, controlled by the states and communist parties. The common features of most migration movements of the socialist countries were the East-to-West direction of most of the long-term population flows, very few return migrations, and hardly any international migration within the former Soviet bloc².

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

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 centrally planned migration patterns. Although international migration as such was hardly existent due to very strict movement control, there were significant population movements between the republics of the 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. These 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 for 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'.

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. The migration balance remains negative in Latvia (1,800) and in Estonia (600). Since 1997, only Lithuania's migration balance became positive again and amounts in 1999 to 1,100 persons (*OECD* 2003).

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 geographical mobility of population in the three Baltic countries is rather low. Relatively low international workers' mobility in the three Baltic states when compared to other EU member countries suggests that inter-regional migration within the Baltic Sea region will be at the heart of matters after the EU enlargement. This calls for more research to be devoted investigating potential migration patterns within this region rather than to speculate about a possible East-to-West mass migration.

2.2 Migration policy

In this Section we argue that because CEE accession countries' citizens are excluded from the free movement rights in the old EU member states, the potential for an inter-regional migration among different parts of CEE accession countries, which do not restrict access to their labour markets, is much higher.

The basic rights attached to citizenship 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

³Positive numbers stand for immigration and negative for emigration.

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 apply to citizens throughout EU-25.

The extension of free movement rights to the additional 75 million new citizens was a particularly sensitive, but 'popular' topic of many intensive debates during the accession negotiations (*European Commission* 2004). On the one hand the ambitions and pressure exerted by some of the new member states and on the other hand the anxiety about possible negative effects on the labour market and employment conditions voiced by some of the old member states, had to be balanced. The solution was found in a rather complex 2+3+2 transitional arrangement referring to workers and proposed by the Commission. In the process of negotiations it has been agreed in the same form with all Eastern and Central European new member states and was included in the Accession Treaty signed on 16 April 2003⁴.

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 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 analysis with those EU member states, which do not restrict access to their labour markets.

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)

13 of 25 member states do not restrict access to their labour markets after the EU enlargement. 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*

⁴The 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.

states nationals. Only two Central and Eastern European accession countries, Poland and Hungary, are known to be applying reciprocal measures only with reference to nationals from the old member states (Table 1).

In contrast the 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 countries⁵ (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 current work permit system with permits 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 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).

Yet another work permit scheme is present in *Denmark*. This country applies a system, under which citizens of the CEE accession countries are allowed to obtain a

⁵The 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.

work permit only once they obtain an official residence permit and only for full time employment. 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).

Even stricter national measures are applied by another two countries. *Italy* operates a work permit scheme, which is automatically limited by an already 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*, as intentions expressed prior to the enlargement might indicate, 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 a 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 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, we want to doubt the usefulness of studying impacts of East-to-West migration ignoring 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 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 that since enlargement there are no legal barriers to workers movement, it is reasonable to assume a high migration potential between such regions. This study investigates recent EU enlargement-related changes in migration patterns among selected NUTS II regions in the CEE accession countries.

3 Existing literature

The high relevance of migration topic in an enlarged Europe has generated a large body of empirical literature that attempts to predict the size of possible labour migra-

tion in the years following the accession. In this Section, some of this 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 temporary regulations and administrative restrictions are not considered in these studies. However, we have shown in Section 2.2 that migration is heavily regulated in EU-25. Therefore, this highly sensitive assumption may reduce the prediction ability of these models (*Kielyte & Kancs* 2002).

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). As shown in Section 2.1, however, there are important differences between the conditions of South-to-North migration and migration from the CEE accession countries. First, labour markets were characterised by full employment and shortages of manual workers in the main receiving countries (Belgium, France, Germany and Switzerland) before the first oil price shock in 1973. Today, on the opposite, unemployment rates are still 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. Bearing these differences in mind, we may conclude that an extrapolation of the South-to-North migration experience is reliable in relative terms, but can be hardly used to estimate migration levels.

At the stage of empirical implementation, 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.

3.2 Methodological issues

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*'. Moreover, as shown in the following, reduced-form models have several limiting assumptions, which are hard to justify.

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 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 relatively precise notion about the development of these conditioning variables in the future. Unfortunately, and in contrast to developed economies, it is 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.

A further implicit assumption in most of these models is that any migration decision is permanent. This is, however, far from being realistic. Migration in CEE is largely a temporary phenomenon, which means that the proportion of the population that will move to another country and perhaps return within a certain period of time is much higher than the proportion that will permanently migrate to a foreign country. There is plenty of evidence showing that, in fact, many East-to-West migrations are actually temporary (see *Dustmann* 1996 for evidence in Europe). It is highly probable that many of the expected future migrations from EU enlargement countries will likewise be temporary. Neglecting this issue may lead to upwards biased estimates of the future migration potential, especially when aggregating up migration flows to obtain stock data at any one point in time.

Last but not least are problems related with 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.

Bearing in mind the described limitations of existing studies we are able to derive properties of a method we would like to use: (i) ability to consider labour market policies; (ii) migration coefficients are estimated from past data of countries under study; (iii) the model allows for country-specific parameters; (iv) endogenous migration rate as well as explanatory variables; (v) allows for a reverse migration; (vi) avoids reverse causality problem among migration rate and explanatory variables.

4 Theoretical framework

Unfortunately, there is no such a model, which would be able to cope with all these described problems. Therefore, in order to overcome shortcomings of the reduced-form models, we suggest to combine two different rather approaches - a NEG model with structural econometrics. For this purpose take a canonical economic geography model derived from *Pflüger* (2004), 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. Econometric estimations are required solely to satisfy properties (ii) and (iii).

4.1 The NEG model: a non-technical overview

This Section spells out the R -region version of *Pflüger's* (2004) 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: $H = H_1 + H_2 + \dots + H_r + \dots + H_R$ with $r \in \{1, \dots, d, \dots, o, \dots, r, \dots, R\}$. 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: $p_{A1} = p_{A2} = \dots = p_{Ar\dots} = p_{AR}$. 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 $r \in \{1, \dots, d, \dots, o, \dots, r, \dots, R\}$.⁶

4.2 Consumption

All consumers share the same quasi-linear utility function:

$$U = \alpha \ln C_x + C_A \quad \text{with} \quad \alpha > 0 \quad (1)$$

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:

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

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 :

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

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

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 :

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

Individual demand (3) can now be written as:

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

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 manufactured firms, each producing a single product. Hence we obtain the constant mark-up equation for profit maximising firms:

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

where p_o is the price of a variety produced in o .

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

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

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

$$\Pi_r = p_r X_r - X_r - W_r \quad (8)$$

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:

$$X_r = W_r (\sigma - 1) \quad (9)$$

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:

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

Alternatively for region o :

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

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 :

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

where $r \in \{1, \dots, d, \dots, o, \dots, r, \dots, R\}$ 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:

$$\Delta V_{do} = V_d - V_o = \alpha \ln\left(\frac{P_o}{P_d}\right) + (W_d - W_o) \quad (13)$$

Using equations (10) - (12), it is straightforward, that ΔV_{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 ΔV_{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 ΔV_{do} must be zero between any pair of regions (13). Thus, we impose $\Delta V_{do} = 0, \forall r$ for each origin-destination pair instead of solving for endogenous values of V_r . The benefits from doing so are twofold: (1) guarantee that the achieved solution is indeed a stable equilibrium ($\Delta V_{do} = 0 \Rightarrow$ 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 here⁸.

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

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

⁸Detailed results are available upon request from the author.

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:

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

where β_i is a conforming vector of utility coefficients, which may vary depending on the identity of the worker⁹, 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 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

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

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:

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

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 :

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

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

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

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:

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

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

$$\ln M_{od} = \ln \left[\frac{n_{od}}{n_{oo}} \right] = \beta_i \Delta V_{i,do} \quad (19)$$

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 model's parameters. It is worth to mention that the approach we use in estimating model's parameters is 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 in the obtained results. 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 Econometric specification

As explained in the introduction, we want to derive an econometrically estimable equation, which contains key parameters of the NEG model. The theoretical framework 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 equation, we use the definition of ΔV_{do} and substitute equation (13) into (19). We obtain the following linear migration equation, where ϵ_{od} is a random prediction error:

$$\ln M_{do} = \beta \ln P_d - \beta \ln P_o + \beta W_o - \beta W_d + \epsilon_{od} \quad (20)$$

Equation (20) is a reduced-form model of migration rate, because we do not observe a NEG model consistent manufacturing price index. Moreover, it does not contain all key parameters of the NEG model. In fact, it contains only the consumers preference parameter α . Thus, (20) needs further transformations before we can estimate it structurally. We use (9) and substitute out W_r ($W_r = \frac{1}{\sigma-1}X_r$ ($\sigma \neq 1$)). In the same time this substitution allows us to obtain a coefficient of σ . To avoid an endogeneity problem, we assume that migration choices at date t are determined from a comparison of V_r across regions at date $t - 1$. Thus, we obtain the following estimable migration equation:

$$\ln M_{od,t} = \beta_1 + \beta_2 \ln P_{d,t-1} - \beta_3 \ln P_{o,t-1} + \beta_4 X_{o,t-1} - \beta_5 X_{d,t-1} + \beta_6 \epsilon_{or,t} \quad (21)$$

where $\beta_2 = \alpha_d$ and $\beta_3 = \alpha_o$, $\beta_4 = \frac{1}{\sigma_o-1}$ and $\beta_5 = \frac{1}{\sigma_r-1}$. This specification provides estimates of key parameters of the NEG framework (α and σ). α is a parameter, which determines consumers' preference for manufactured goods and σ is the elasticity of substitution. Note that our estimable equation (21) is very similar to a canonical reduced-form model. In fact (21) is a reduced-form model, if applied directly for predicting migration flows.

There are several issues to address before performing estimations. A major difficulty with equation (21) arises from the definition of the traditional sector. According to the theoretical framework, the 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 stands for all manufactured goods.

Another difficulty with equation (21) arises from the definition of the immobile factor. According to the theoretical framework, the difference between the L ('immobile factor') and H (labour) lies in the inter-regional mobility - while regional supply of the former is fixed, labour is mobile between regions. Second difference lies in the usage of the two factors - traditional sector uses only the immobile factor (L), while manufacturing uses labour (H) as a fixed cost and the immobile factor for the variable cost. Unfortunately, we do not have detailed production 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 Data

Data requirements of the NEG model and econometric estimations are fundamentally different. Therefore, two data sets are required - one for estimating coefficients of equation (21) and one for calculating the NEG model.

The estimation of equation (21) requires panel data of migration rate M_{od} , and of all explanatory variables (H , P 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 *Regio* data covers 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. 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 reasons, we do not pool all the data together but perform separate regressions for each country. To fill gaps in the data and make the data internally consistent, we are reliant on national statistics' data, which we use extensively. Thus, the information contained in *Regio* together with the national statistics allows us to obtain bilateral migration data for 14 years building a panel of $14 \times 3 = 42$ observations. 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 price index); and (2) the time coverage in the national statistics differ among countries substantially. 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 the national statistics allows us to obtain data for H_r , L_r , P_r and X_r for each NUTS II region in the three Baltic states building equally sized panels with 42 observations.

The NEG model is considerably less data demanding and requires a cross-Section of NUTS II regions of one year only. In particular, we need values for H_r , L_r and models parameters. Rest of the variables are calculated endogenously within the model. For this purpose we use data already prepared for the estimation of equation (21). 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.

5.3 Estimation results

We perform estimations of equation (21) for three CEE accession countries: Estonia, Latvia and Lithuania. There are several reasons for selecting these three out of ten CEE accession countries. First, Baltic states, due to the fact they were part of the former Soviet Union, have a common migration history. The language and cultural differences between the three countries are rather low (Section 2.1). Second, Baltic states as any other of the new EU member states do not restrict access to their labour markets by new EU member states nationals (Section 2.2). Because of these two reasons, we expect that after opening labour markets main migration flows will be among similar CEE accession countries instead of commonly assumed East-to-West¹⁰ migration. Last but not least, the data set we compiled covers only the three Baltic countries¹¹.

Thus, we use the historical experience of Estonia, Latvia and Lithuania, since in line with existing studies we have to assume that the same households will respond to the same factors in the same way in the future. Equation (21) is estimated by ordinary least squares. Table 3 reports our estimation results¹².

Table 3. Dependent variable: log migration rate

	$\log P_{d,t-1}$	$\log P_{o,t-1}$	$X_{o,t-1}$	$X_{d,t-1}$	R^2
R_{12}	-0.331 (0.093**)	0.315 (0.102**)	-0.214 (0.041**)	0.240 (0.021**)	0.86
R_{13}	-0.357 (0.083**)	0.359 (0.090**)	-0.251 (0.107*)	0.262 (0.024**)	0.88
R_{23}	-0.326 (0.077**)	0.328 (0.084**)	-0.271 (0.111*)	0.203 (0.090*)	0.91

OLS estimates, with standard errors presented under estimated coefficients. $R_1 = Estonia$, $R_2 = Latvia$, $R_3 = Lithuania$. All variables lagged by one year. *significant at 5%, **significant at 1%.

Most parameters, in particular those which are presented in Table 3, converge toward consistent values. Although the estimation results do show substantial differences among the three Baltic countries, the coefficients always have the same order of magnitude. All estimated coefficients have the expected signs. Destination region's

¹⁰Most of the empirical studies available in the migration literature (Section 3.1) have a strong East-West focus. Although, a politically appreciated experiment of thoughts, because of restrictive migration policies, it is of limited practical relevance. Furthermore, Section 3.2 has demonstrated why without explicitly controlling for temporary regulations and administrative labour market restrictions, the obtained results are biased.

¹¹A work towards extending the data set to all ten CEE accession countries has already been taken up.

¹²For convenience, only NEG-model relevant coefficients are reported here. All results are available upon request.

price index, $\log P_{d,t-1}$, is inversely related to the migration rate - as higher are manufacturing prices in the destination region, as lower utility and lower incentives to migrate to the particular region. The opposite is true for the origin region's price index, $\log P_{o,t-1}$, - as higher manufacturing price index, as lower utility and as higher incentive for mobile workers to leave the region. The positive sign is in line with our expectations and holds true for all three Baltic countries. The relationship between industrial output, $X_{r,t-1}$, and migration rate is inverse to the industrial price index. A higher level of industrial output implies a higher level of utility, and hence according to equation (13), mobile workers have less incentive to leave the region. Therefore, the estimated coefficients are negative. Analogously, the opposite holds for industrial output in the destination region, $X_{d,t-1}$. These coefficients are positive and are in line with our expectations (column $X_{r,t-1}$ in Table 3).

Regarding the magnitude of numerical estimates, our results are in line with existing studies. Coefficient β_3 implies that an increase in this year's manufacturing price index in the host region, $P_{o,t-1}$, will ceteris paribus result in a next year's increase of net migration rate into region d by 0.315 to 0.359 percentage points. With a relatively high value of β_2 and β_3 , and the fact that the coefficients are significant, manufacturing price index turn out to be the most important independent variable in this estimate. Besides the fact that different specifications are applied and different countries studied, most of the previous empirical studies point to a certain dimension of potential migration following the EU enlargement (*Bauer & Zimmermann 1999, Fertig 2000, Boeri & Brücker 2001, Kielyte & Kanacs 2002*). The overall statement is primarily 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 3-5 % of the CEE-10's current population is expected to migrate to EU-15 countries in the medium and long run (10-30 years). In the case of the 5 % estimate proving accurate the actual number of migrants would correspond to the present population of Denmark, or somewhat more than half of the Swedish population. However, more modest migration will only reach 0.8-1.0 % of the present EU-15 population. Deviations in our estimates might be caused by statistical problems, small number of observations being one of them. Other potential problems current results might be associated with are heteroscedasticity, or temporal autocorrelation.

5.4 Transport costs

Last parameter the NEG model depends on 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 NEG models (*Head & Mayer 2004*) and assume symmetric bilateral barriers $\phi_{od} = \phi_{do}$. This assumption leads to a very simple estimator for ϕ_{od} :

$$\hat{\phi}_{od} = \sqrt{\frac{m_{od}m_{do}}{m_{oo}m_{dd}}} \quad (22)$$

where m_{od} is import of goods and services from origin region o to an destination region d . The numerator in equation (22) requires only trade flow data expressed according to 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 index of transport costs, which ranges from zero to one. Table 4 reports the estimation results.

Table 4. Road distances and transport cost

<i>Regions</i>	D_{od} (km)	T_{od}
R_{12}	297.1	0.969
R_{13}	608.6	7.764
R_{23}	290.8	2.390

D_{od} - road distances between capitals (*Route 66*)

Source: Own calculations based on *Comext* (2004).

Table 4 reports values of T_{od} , which are calculated using the definition of $\hat{\phi}_{od}$ ($\phi_{od} = T_{od}^{1-\sigma}$). $\hat{\phi}_{od}$ are estimates obtained from estimating equation (22). 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 (T_{13}) are considerably higher than trade costs among any other two regions¹³. These remarkable differences in general 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 regions world (Estonia, Latvia, and Lithuania), transportation of goods between Estonia and Lithuania has not only the biggest average transport distances, in addition all goods always have to cross two borders¹⁴. Border-crossing costs are part of the general transport costs T_{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 two peripheral border regions faster than among core and peripheral regions. Practically this means that T_{13} should decrease faster than T_{od} estimates for any other pair of regions.

¹³ 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.

¹⁴Bilateral trade between any other pair of regions involves crossing only one border.

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 from Section 4, the estimated parameters from Section 5, and *Regio* data from 2003, we are able to calculate a long-run equilibrium solution for different levels of trade costs, which represent different levels of EU integration.

6.1 Base run

We use a cross-section *Regio* data of 2003 for each of the three Baltic states to implement and run the NEG model. This equilibrium solution 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 as a share of regions' stock with mobile workers¹⁵.

We start with solving the model for a long run equilibrium, which allows us to obtain values of all endogenous variables, such as prices, output, wages, sectoral employment and bilateral migration flows for each region (if any). We use the model developed in Section 4, fill it with the *Regio* data and the estimated parameters' values. We obtain endogenous values of all endogenous variables, two of which are reported in Table 5¹⁶. For convenience, this solution is called 'base run'.

Table 5. Base run¹⁷

<i>Regions</i>	P_r	W_r
R_1	1.4704	0.1362
R_2	1.4705	0.1361
R_3	1.5224	0.1479

For assumptions see Section 4.

Source: NEG model simulations.

Although not reported here, the indirect utility is equal in all three regions in the 'base run' and the utility differences (ΔV_{do}) are zero between any two pairs of regions. Hence, in the 'base run' there is no incentive to migrate and migration rates are zero between any two pairs of regions (column M_{od}^{BR} in Table 6). Zero migration rates in the light of different equilibrium values for P_r and W_r require a more detailed

¹⁵In the following this share is referred to as a migration rate.

¹⁶The model is implemented by the General Algebraic Modelling System (GAMS) and solved using CONOPT and MINOS (*Brooke et al 1988*).

¹⁷For convenience, only migration-scenarios relevant variables are reported here. All results are available upon request.

explanation. Indeed, we might expect that the model would predict inter-regional migration if the price index and wage rate is different among the regions. The key parameter giving answer to this question is inter-regional transport cost (T_{od}), which also determines the indirect utility. As reported in Table 4, they are highly different among the three regions. Multiplying the trade volume of each partner country with the respective transport costs, we obtain a volume of transport costs. Transport costs combined with the producer price (equation 4) provide estimate of the regional price index (P_r), which is higher in Lithuania, although, it has the highest wage rate (column P_r in Table 5).

Further experimenting with the model, if trade costs would be equal among all regions (combined with different regions' sizes in terms of factor endowment), the model reports differences in the indirect utility among regions and, hence, a positive inter-regional migration. If trade costs would be equal among all regions combined with equal regions' sizes in terms of factor endowment, the model predicts equal values for price index, wages and indirect utility in all regions. Hence, there is no incentive for migration.

6.2 Simulation exercises

We follow *Faini, Melo & Zimmermann* (1999) 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 for transportation cost changes.

The real transport costs at benchmark have already been estimated in Section 5.4. Reliable estimates on transportation cost changes related to the European integration are not available in the literature. To overcome this deficiency, we construct several hypothetical scenarios, where transport costs between the peripheral border regions (T_{13})¹⁸ are reduced by 10%, 20%, 30%, 40%, 50%, 60% respectively. Although, we have no particular reason to believe that European integration will induce transport costs reduction in this order of magnitude or in such an asymmetric way favouring the peripheral regions, the 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 long-run equilibrium solution, because these differences are an incentive for relocating. Since we are interested in a long-run

¹⁸Transport 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.

equilibrium solution, 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 $\Delta V_{do} = 0$ for all regions, and solve the model for \hat{H}_r . Migration rate is then calculated as $M_r = H_r - \hat{H}_r$ ¹⁹.

Table 6 reports simulation results, where $T = 100\%$, $T = 80\%$, $T = 60\%$, and $T = 40\%$ are the respective values of transport costs. Columns 2-5 report the estimated migration flows as a percentage of employed home population. Considering estimates in Table 6 it is straightforward to identify that a different level of transport costs lead to globally consistent²⁰ 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 regional integration follows a pattern we assumed in the simulations. As already explained, initially there is no migration (column M_{od}^{BR} in Table 6). When regions integrate, Lithuania seems to be the biggest gainer from trade cost reduction. It steadily gains workers throughout the simulations (row R_3 in Table 6). The immigration rate in Lithuania is continuously increasing from 7.076% when transport costs with Estonia are reduced by 20% to 7.250% when T_{13} is reduced by 60%.

Table 6. Transport costs and net migration rate²¹

	$M_r^{BR}(\%)$	$M_r^{20}(\%)$	$M_r^{40}(\%)$	$M_r^{60}(\%)$
Regions [†]	$T_{13} = 100\%$	$T_{13} = 80\%$	$T_{13} = 60\%$	$T_{13} = 40\%$
R_1	0.000	-11.025	-9.069	1.906
R_2	0.000	-3.911	-5.112	-11.755
R_3	0.000	7.076	7.112	7.250

[†] $R_1 = Estonia$, $R_2 = Latvia$ and $R_3 = Lithuania$.

See Sections 5 and 6.2 for assumptions.

Latvia seems to be the largest loser from the integration process with the EU, when transport costs are reduced asymmetrically favouring the peripheral regions. The emigration rate in Latvia is continuously increasing from 3.911% when transport costs are reduced by 20% to 11.755% when T_{13} is reduced by 60% (row R_2 in Table 6).

Estonia is probably the most interesting region from an analytical perspective and might be subject to predictions of the NEG theory - there is a non-monotonic

¹⁹Analytically 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.

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

²¹Positive migration rate = immigration; negative migration rate = emigration.

relationship between the level of regional integration and the regions' share of mobile factors. 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 integration the smaller region, R_1 loses 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 gain the share of mobile workers (row R_1 in Table 6).

6.3 Gross migration flows

In the previous Section we obtained a net migration rate for each region. The net migration rate, however, does not provide information about the regional breakdown of emigrants' host regions and immigrants' source regions. In Subsection 4.5 we derived an expression (19), which enables to calculate bilateral migration flows. In this Section we have to implement equation (19) into the NEG model to simulate bilateral region-to-region migration flows. Before we can calculate bilateral region-to-region migration flows, the NEG model requires some minor adjustments. Particularly, the endogenous variable of interest now is M_{od} , and the required equation is (19). The remaining equilibrium expressions (10), (11), (13) derived in Section 4.4 do not change.

The endogenous variables we are interested in this Section are P_r , W_r , \hat{H}_r , M_{od} . Hence, the simultaneous system of equations we are going to solve in this Section contains $(3 \times R) + (\frac{R}{2} \times \frac{R}{2})$ unknowns and $(3 \times R) + (\frac{R}{2} \times \frac{R}{2})$ equations²². $3 \times R$ endogenous variables are P_r , W_r , \hat{H}_r and $\frac{R}{2} \times \frac{R}{2}$ endogenous variables are bilateral migration flows, M_{od} , for each pair of regions. $3 \times R$ equations are (10), (11), (13) and $\frac{R}{2} \times \frac{R}{2}$ equations are (19) for each pair of regions. Thus, size of the model is much larger, when bilateral gross migration flows are calculated. The number of equations and the number of endogenous variables increases by $[\frac{R}{2}]^2$.

Table 7 reports simulation results, where $T = 100\%$, $T = 80\%$, $T = 60\%$, and $T = 40\%$ are the respective shares of initial transport costs. Columns 2-5 report the estimated bilateral migration flows for each pair of regions. Estimates in Table 7 are not comparable straightforwardly to the estimates in Table 6, because figures in Table 7 report numbers of movers and figures in Table 6 report percentage changes in migration rate compared to the Base Run. Nevertheless, estimates reported in Table 7 are in line with those presented in the previous Section. Therefore, interpretation of the results is not repeated here.

Why these results are different from those one obtains estimating a reduced-form

²²Generally, it is possible to calculate $R \times R$ migration flows. However, the way the theoretical NEG model is constructed (the only driving force of migration are inter-regional differences in indirect utility) allows to calculate only $\frac{R}{2} \times \frac{R}{2}$ migration flows implicitly assuming that there is no two-way migration between any two regions.

model, such as (21)? Looking inside the 'black box' (NEG model) we recognise that in the model changes in transport costs are absorbed through changes in other variables (e.g. prices, wages, quantities produced, consumed and quantities transported), until a new equilibrium is reached. Because of the changes in these other 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 workers immigration (firms' entry)²³. 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 through 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 international 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.

Table 7. Transport costs and gross migration flows²⁴

	M_{od}^{BR}	M_{od}^{20}	M_{od}^{40}	M_{od}^{60}
Regions [†]	$T_{13} = 100\%$	$T_{13} = 80\%$	$T_{13} = 60\%$	$T_{13} = 40\%$
R_{12}	0.000	0.000	0.000	-1123
R_{23}	0.000	3962	5132	10677
R_{31}	0.000	-6495	-5343	0.000

[†] $R_1 = Estonia$, $R_2 = Latvia$ and $R_3 = Lithuania$.

See Sections 4.5, 5 and 6.2 for assumptions.

These induced effects, which work in a reverse direction (from left-hand-side towards right-hand side variables), are not captured in a standard reduced-type approach. Without considering changes in explanatory variables, induced by workers migration and industry relocation, usually would result in biased results. Moreover,

²³Note 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 $H_r = N_r$.

²⁴Positive migration rate \Rightarrow migration from region o to region d ; negative migration rate \Rightarrow migration from region d to region o .

because of a reverse causality, reduced-form models' estimates will likely be biased. The NEG model applied in this paper captures all changes in explanatory variables. The ability to capture these effects and to avoid problems caused by reverse causality in a theoretically consistent and a data-conform way is the main innovation of the current paper.

7 Conclusions

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 European integration. Model's parameters are estimated using a semi-structural equation, which is derived entirely from the theoretical NEG model.

Although, the New Economic Geography has a history of almost two decades, its usage in migration literature is still rather limited.²⁵ 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 critiques 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 - a certain reduced-form specification might still perform better in terms of forecasting performance. Therefore, we urge for more research, both methodological and empirical, be devoted to estimating structural NEG models in predicting location of firms and workers. Future expectations also play a significant part in a migration decision. 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 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 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 peripheral at the beginning of integration, turns out to be the biggest winner in terms of industry and mobile workers.

²⁵ *Crozet* (2004) is one of the notable exemptions.

References

- [1] Baldwin R., R. Forslid , Ph. Martin, G. Ottaviano and F. Robert-Nicoud. (2003), *Economic geography and public policy*. Princeton, N.J., Princeton University Press.
- [2] Bartel, Ann (1989), Where Do the New U.S. Immigrants Live? *Journal of Labor Economics*, 7, 371-391.
- [3] Bauer, Thomas, and Zimmermann, Klaus F., (1999), *Occupational Mobility of Ethnic Migrants*, Discussion Paper No. 58, IZA, Bonn, Germany, September.
- [4] 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.
- [5] Brooke, A., David Kendrick and Alexander Meeraus (1988), *GAMS: a User's Guide*, New York: The Scientific Press.
- [6] Commission of the European Communities (2004), *First Annual Report on Migration and Integration*, European Commission, Brussels.
- [7] Crozet, M. (2004), Do migrants follow market potentials?, *Journal of Economic Geography*. 4, no. 4, pp. 439-458.
- [8] 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.
- [9] 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.
- [10] Dustmann, C. (1996), Return Migration: The European Experience, *Economic Policy*, 22, pp.213-242.
- [11] Eurostat (2004), *Central European countries' employment and labour market review*, Luxembourg.
- [12] Faini R., de Melo J., Zimmermann KF (1999), *Migration: The Controversies and the Evidence*, London: Cambridge University Press for CEPR.
- [13] 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.

- [14] 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.
- [15] 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.
- [16] Harris, J.R., M.P. Todaro (1970), Migration, Unemployment and Development, A Two-Sector Analysis, *American Economic Review*, 60, pp. 126-142.
- [17] Hatton TJ (1995), A Model of UK Emigration, 1870-1913. *Review of Economics and Statistics*, 77:407-15.
- [18] Head, Keith and Thierry Mayer (2004), The Empirics of Agglomeration and Trade, in *Handbook of Regional and Urban Economics*, vol. 4.
- [19] Kielyte, J., Kanacs, A. (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.
- [20] Krugman, P. (1991), ‘Increasing Returns and Economic Geography’, *Journal of Political Economy*, 99, 3, 483–99.
- [21] Krugman, P., A. Venables (1995), Globalization and the Inequality of Nations, *Quarterly Journal of Economics*, 110, 857-880.
- [22] Levine, Phillip and David Zimmerman (1999), An Empirical Analysis of the Welfare Magnet Debate Using the NLSY, *Journal of Population Economics*, 12, 391-409.
- [23] McFadden, D. (1973), Conditional Logit Analysis of Qualitative Choice Behavior, in P. Zarembka, ed., *Frontiers in Econometrics* (Academic Press, New York).
- [24] OECD (2003), *Labour Market and Social Policies in the Baltic Countries*, Paris.
- [25] Pflüger, M. (2004), ‘A Simple, Analytically Solvable, Chamberlinian Agglomeration Model’, *Regional Science and Urban Economics*.
- [26] Sjaastad L.A. (1962), The Costs and Returns of Human Migration, *Journal of Political Economy*, 70:80-93.
- [27] Straubhaar T., Zimmermann K.F. (1993), Towards a European Migration Policy, *Population Research and Policy Review*, 12:225-41.