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

D'Artis Kancs [1]

Abstract

In this paper I 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 my approach is that it endogenises both, explanatory variables and the migration rate. Model's parameters are estimated econometrically using a structural equation, which is derived entirely from the theoretical NEG model. My 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, Agglomeration.

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). 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 levels.

---

[1] d'Artis Kancs, London School of Economics, London, UK. Email: kancs@lse.ac.uk.

***M**igration **L**etters*, Volume: 2, No: 1, pp. 32– 63 April 2005
(ISSN: *print:* 1741-8984 & *online:* 1741-8992)

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 models and, in addition, focus on East-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. The current paper develops an alternative methodology - a structurally estimable NEG model of migration flows and provides a geographically relevant application of the model.

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 migration function, since 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 simulation to the period after CEE integration with the EU, when free movement of workers is introduced. European integration is modelled by altering model's parameters - reducing general transport costs.

Section 2 gives an overview about historical migration patterns in the new member states. After providing an overview of the historical migration trends for selected accession countries of Central and Eastern Europe, section 3 critically assesses the existing literature and discusses methodological shortcomings of theses 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 comparison with existing studies are presented in section 6.

2 Migrations in Central and Eastern Europe

This Section focuses on the main past and recent patterns of international migration concerning the new EU members and the accession countries, as well as on the identification of the migration policies, push and pull factors influencing the observed migration phenomena. Among others, the relationships between migration and the economic and policy factors are discussed. The analysis is followed by the identification of major directions of the future flows, as well as of the main origin and destination countries.

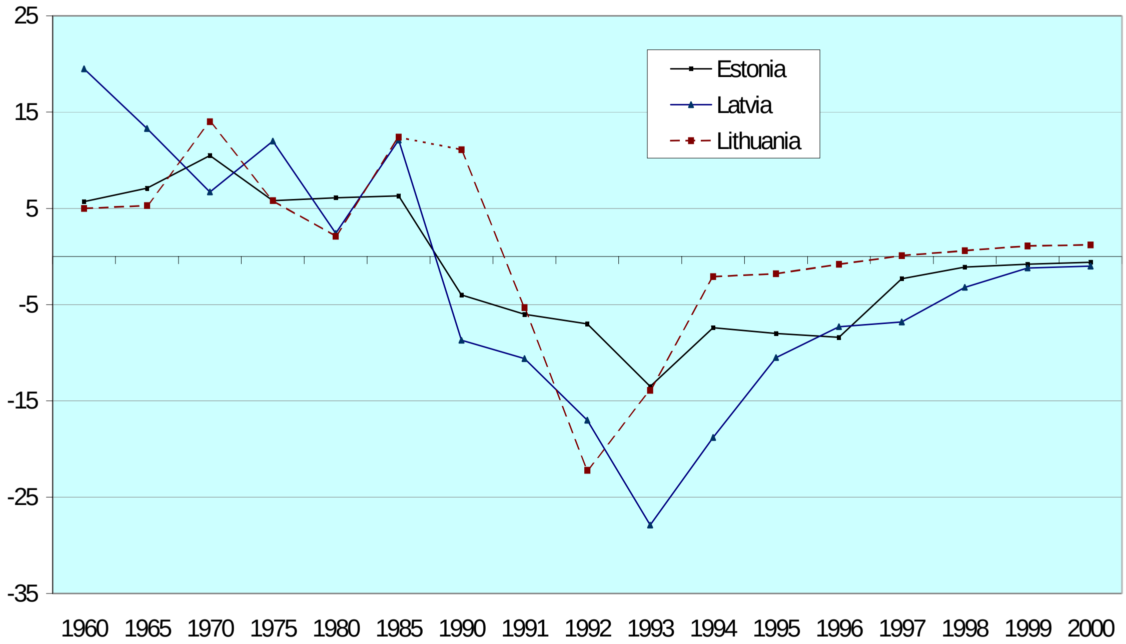
2.1 Historical overview

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 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 inelastic labour market, which was supported by imported foreign labour force. Only the oil crisis of 1973 resulted in change in migration policies, but at that time the momentum of immigration was high and Western Europe was becoming multiethnic and multi-cultural.

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 migration within the former Soviet bloc[2] (apart from the countries of the former Soviet Union), mainly because of strict movement control. Baltic States, due to the fact they were part of the former Soviet Union, had different migration experiences (*Kielyte* & *Kancs* 2002).

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



Source: Eurostat (2004) and Kielyte & Kancs *(2002).*

The 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 (see Figure), comprised mainly of the immigration of Russians, many of whom were the Soviet military personnel. These phenomena contributed to significant changes of the ethnic structures of all three Baltic republics (*Kielyte* & *Kancs* 2002).

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

After more than 30 years, net migration became negative for all three countries for the first time in 1990 (see Figure)[3]. 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 outmigration of the so-called Russian speaking population (*Kielyte* & *Kancs* 2002).

Data for 2000s indicate the most recent migration trends among the 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).

2.2 Migration policy

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 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 negotia-

---

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

tions 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 2003[4].

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 possible)*
Estonia	--	Slovakia	--
Hungary	2 (+3 +2 possible)*	Slovenia	--
Ireland	--	Sweden	--
Latvia	--	UK	Workers' 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 member states that decided to open up their labour markets as from the first day of the accession (Table 1). Most new member states also opened up their labour markets (Table 1).

[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.

Out of the 8 Central and Eastern European new member states only Poland and Hungary are known to be applying reciprocal measures with reference to nationals of the old member states (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.*

In contrast the majority of the old 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 restricting access to their labour markets

Country	**Transitional measures on labour movement (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 per year). 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

[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.

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 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 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, one might doubt the usefulness of studying impacts of East-West migration ignoring East-East migration. In contrast, 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 the welfare differences are increasing among different regions. Bearing in mind that per capita income in some NUTS 2 regions is twice as high as in other regions in the CEE accession countries and that 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 among selected NUTS 2 regions in the CEE accession countries.

3 Existing literature

The high relevance of the topic 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, 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 studies

Studies that forecast future migration flows from the accession countries to the EU are typically based on econometric models that using historical data on migration flows from countries other than the CEECs (see *Fertig* 2000, *Boeri* and *Brücker* 2001, *Bauer* and *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.

Besides the fact that different specifications are applied, all of the studies point to a certain dimension of potential East-to-West migration. 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.

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 by these studies. However, I will argue that migration is heavily regulated in EU-25. Therefore this incorrect assumption may reduce the prediction ability of these models.

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* and *Zimmermann* 1999). There are, however, important differences between the conditions of South-North migration and migration from CEE new member states. 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 of East-West migration 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 these CEECs than in traditional sending countries. We may conclude that extrapolation of the South-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* and *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 Estonia. 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 (regarded by the authors themselves), 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* and *Brücker* 2001, and *Fertig* 2000) refer to *Hatton's* (1995) more elaborate model as a motivation for their 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 another 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 devel-*

oped here is only one among many different specifications that could be developed'. Moreover, as shown in the following, econometric specification and methodology are in themselves highly critical.

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 behavioural relationships to hold or one has to have a relatively precise notion about the development of these conditioning variables in the future. This of course is a difficult problem in itself, and adds substantial error to migration predictions. 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 suitable to estimate 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 live in a foreign country at a given point in time. There plenty of evidence showing that in fact many East-West migrations are actually temporary (see *Dustmann* 1996 for evidence in Europe). It is likely 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 the ad hoc way of specifying the migration equations. Depending on assumptions about the distribution of the costs of migration and preferences across the population, one can conceive different functional forms. The fact that results are highly dependant upon the specification used underlines once again that the obtained results are highly sensitive to the particular specification. Because of these problems, the reduced-form models are reliable in relative terms, but are biased when used to estimate migration levels.

I cannot list all shortcomings of the traditional reduced-form econometric approach, but I hope these will show that how much caution one should have, when considering these results. To be clear I have grouped these deficiencies into two groups: empirical and methodological. The main *empirical shortcomings* we identified are: (i) assumption of open labour markets in the receiving countries; (ii) problems associated with extrapolation of North-South migration experience; and (iii) assumptions implied at the stage of empirical implementation (e.g. slope (response) parameters are the same for all countries). The main *methodological shortcomings* we identified are: (iv) bias associated with reduced form models, where explanatory variables are kept exogenous to the model; (v) assumption that all migration flows are of permanent character; and (vi) problems related with the ad hoc way of specifying the estimation equations.

4 Theoretical framework

In order to overcome shortcomings of the reduced-form models, one needs to relay on a rather structural approach. For this purpose we develop a canonical economic geography model derived from *Pflueger* (2004), which in turn represents an analytically solvable version of *Krugman's* (1991) core-periphery model[6]. In that model, we are able to study

[6] Krugman (1991) shows 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.

the impact of increasing regional integration on the internal geography (location of workers and firms) of regions. By adopting the NEG model as a theoretical framework, we are able to cope with the critiques (ii)-(iv) mentioned above. The critique (ii) has been addressed by estimating parameters from the same regions' historical data we are predicting migration, i.e. Baltic regions. The critique (iii) has been addressed by using a NEG model, which takes into account each region separately; therefore, we could obtain different slope parameters for each region. The critique (iv) has been addressed by using a general equilibrium model, which takes into account adjustments in explanatory variables.

4.1 The NEG model: a non-technical overview

This Section spells out the R -region version of *Pflueger's* (2004) geography model. The 'world' consists of R regions, endowed with two factors, one immobile factor (L) and 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 et 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', the perfectly competitive immobile sector, will serve as 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. Finally, we use the traditional good as a numeraire, therefore 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 $r \in \{1,..,d,..,o,..,r,..,R\}$[7].

4.2 Consumption

All consumers share the same quasi-linear utility function:

(1) $U = \alpha \ln C_x + C_A$ with $\alpha > 0$

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:

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

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 be-

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

tween two varieties ($\sigma > 1$). Given income Y, each consumer maximises 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:

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

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 $T_{oo} = 1$. Using (2) and (3) we are able to derive the following industrial price index for each region d:

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

Individual demand (3) can now be written as:

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

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 $TC_r\left(x_j\right)=W_rH_j+L_rx_j$, where W_r represents the compensation of labour in region r. Hence, $TC_r\left(x_j\right)$ 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:

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

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:

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

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

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

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:

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

4.4 Equilibrium conditions

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

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

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 :

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

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

In the long run, workers are mobile between regions. They migrate towards regions with the highest indirect utility. By migrating between regions workers equalise price indices, real wages and, hence, indirect utilities among regions. 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:

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

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

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

Using equations (10) - (12), it is straightforward, that (ΔV_{od}) only depends on the share of workers in region d and the parameters of the model.

5 Empirical implementation

Data requirements of the NEG model and econometric estimation are fundamentally different, the two data sets are described separately - one for estimating the coefficients of equation (15) and one for calculating the NEG model.

The estimation of equation (15) requires panel data of all explanatory variables (H, P and X). *Regio* database (*Eurostat* 2004) provides annual bilateral migration data at a regional level. While invaluable, this information has three main shortcomings. First, data are limited to intra-country migration since data on migration from a region in one country to a region in another are not available. Second, *Regio* does not provide data at a very detailed geographic level; data are available at the NUTS 2 or NUTS 3, depending on the year and country. Finally, the definition of migration used and the time span covered by the data is not exactly the same for all three Baltic countries. For these reasons, we do not pool all data together but perform separate regressions for each country.

In addition to *Regio* data, national statistics provide data on employment, regional price index and industry output at a regional level. From the national statistics we are able to extract an internally consistent information for 14 years (1989 - 2003) for 20 NUTS 3 regions in the Baltic States (5 for each Latvia and Estonia, and 10 for Lithuania). 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 3 region in the three Baltic States, and our panel contains 70 observations for each Latvia and Estonia and 140 observations for Lithuania.

The second data set is required for the NEG model, which is considerably less data demanding and requires a cross-section of NUTS 2 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. This data is readily available from *Regio* for the year 2003. The parameter values are obtained, as already mentioned, from separate regressions, which are explained in more detail in the following two subsections.

5.1Econometric specification

We start with a specification, which is directly taken from the theoretical NEG model. Based on *McFadden* (1973), if and only if the R disturbances $\varepsilon_{i,od}$ 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. Substituting equation (13) this conditional logit function, we obtain the following non-linear equation:

(14) $$\ln M_{do} = \alpha \ln \frac{P_d}{P_o} + W_o - W_d - \ln g$$ [8]

Because no statistical data is available on wages at a regional level in the Baltic countries, we use equation (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 $t-1$. Introducing β into equation (14), we obtain the following linear estimable equation:

(15)

$$\ln M_{do} = \beta_1 + \beta_2 \ln P_{d,t-1} - \beta_3 \ln P_{o,t-1} + \beta_4 X_{o,t-1} - \beta_5 X_{d,t-1} + \beta_6 u_t$$

where $\beta_2 = \alpha_d$ and $\beta_3 = \alpha_o$, $\beta_4 = \frac{1}{\sigma_o - 1}$ and $\beta_5 = \frac{1}{\sigma_d - 1}$. 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 estima-

---

[8] g captures rest of the equation and is equal to $^{R}_{r\;1}\, e^{\;V_{i,}}$. In a three regions example it can be written as ($^{R}_{r\,1} e^{V_{i,od}} \left(\frac{P_2}{P_1}\right) e^{W_1\,W_2} \left(\frac{P_3}{P_1}\right) e^{W_1\,W_3}$).

ble equation (15) is very similar to a canonical reduced-form model. In fact, it 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 (15) 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 market structure and the presence of scale economies: the `traditional' sector should stand for 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 (15) 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').

5.3 Estimation results

I perform estimations of equation (15) for three CEE accession countries: Estonia, Latvia and Lithuania. There are several reasons for selecting the three 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. Second, Baltic States, as any other of the new member states do not restrict access to their labour markets by new EU member states nationals. Third, because of (1) and (2) I expect that main migration flows will be among similar CEE accession countries instead of commonly assumed East-West migration.

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. We choose panel fixed effects estimation technique on the basis of diagnostic tests and willing to account for country-specific time-invariant factors that influence migration. Equation (15) is estimated by ordinary least squares with a full set of fixed effects. Table 3 reports our estimation results[9].

Table 3. Dependent variable: log migration rate

	Estonia	**Latvia**	**Lithuania**
$\log P_{d,t-1}$	*-0.333*	*-0.359*	*-0.328*
	*(0.094)***	*(0.084)***	*(0.076)***
$\log P_{o,t-1}$	*0.314*	*0.358*	*0.327*
	*(0.101)***	*(0.090)***	*(0.082)***
$X_{o,t-1}$	*-0.215*	*-0.252*	*-0.272*
	*(0.040)***	*(0.106)**	*(0.112)**
$X_{d,t-1}$	*0.241*	*0.264*	*0.204*
	*(0.025)***	*(0.026)***	*(0.091)**
No Obs.	*70*	*70*	*140*
R^r	*0.86*	*0.88*	*0.91*

*OLS estimates, with standard errors presented under estimated coefficients. All variables lagged by one year. * significant at 5%; ** significant at 1%.*

Most parameters, in particular those which define the price index, market and supplier access, converge toward consistent values. Although the estimation results do show substantial differences among the three Baltic countries, the

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

coefficients always have the same order of magnitude. All estimated coefficients have expected signs. Destination region's 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, $\log P_{d,t-1}$, - 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 (row $X_{r,t-1}$ in Table 3).

Generally, these results are in line with existing studies, however they are subject to many statistical problems, small number of observations being one of them. Other potential problems current results might be associated with are heteroscedasticity, or temporal autocorrelation. As the cross sectional units of the panels are NUTS III regions, potentially this might cause problems of spatial 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 and assume symmetric bilateral barriers $\phi_{od} = \phi_{do}$. This assumption leads to a very simple estimator for ϕ_{od}:

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

where m_{od} is import of goods and services from origin region o to an destination region d. The numerator in equation (16) 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

*Regions**	D_{od} *(km)*	T_{od}
R_{12}	*297.1*	*0.969*
R_{13}	*608.6*	*7.764*
R_{23}	*290.8*	*2.390*

**R_1: Estonia, R_2: Latvia, and R_3: Lithuania. 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 $\hat{\phi}_{od}$ (ϕ_{od} = $T_{od}^{1-\sigma}$). $\hat{\phi}_{od}$ are estimates obtained from estimating equation (16). Recalling that $0 < \phi_{od} < 1$ with 0 denoting prohibitive trade costs, the overall level of trade freeness appears to be quite low, even though we have calculated $\hat{\phi}_{od}$ for pairs of countries known for their high levels of formal trade integration (BAFTA - Baltic Free Trade Agreement).

Table 4 reveals that trade costs between Estonia and Lithuania (T_{13}) are considerably higher than trade costs among any other two regions. 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'. In our three regions word (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 borders. Border-crossing costs are part of the general transport costs T_{od} and were comparatively high before integration. Bilateral trade between any other two regions involves crossing of only one border.

6 Predicting migration flows

This Section is 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 regional 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 (Subsection 'Simulation exercises') we then compare the obtained estimates with the 'base run' and calculate differences in the regions' stock with mobile workers.

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 employ the model developed in Section 4, and use the *Regio* data and the estimated parameters' values from Section 5 to solve the model. We obtain endogenous values of all endogenous variables, tow of which are reported in Table 5. For convenience, this solution is called 'base run'.

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 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 costs (T_{od}), which also determines the indirect utility. As reported in Table 4, they are highly different among the three regions. If we multiply the trade volume of each partner country with the respective transport costs, we obtain average transport cost. Transport costs combined with the producer price (equation) 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).

Table 5. Base run

Regions	P_r	W_r
R_1	*1.4764*	*0.1342*
R_2	*1.4765*	*0.1341*
R_3	*1.5294*	*0.1459*

For assumptions see Section 4. Source: Own calculations.

On the opposite, 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* and *Zimmermann* (1999) and model European integration and increased labour mobility as a reduction in transport costs. In order to set up an integration scenario of decreasing transport costs one requires: (1) Magnitude of the real transport costs at benchmark, and (2) 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 (not known to the authors) in the literature. Therefore, we investigate labour market implications in several hypothetical scenarios, where transport costs between Estonia and Lithuania (T_{13}) 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 short-run equilibrium we then obtain differences in the price index, wages and indirect utility between regions. Clearly, this is not a long-run equilibrium solution. We are interested in a long-run 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 $M_r = H_r - \hat{H}_r$.

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 reports the estimated migration flows as a percentage of employed home population.

Table 6. Transport costs and migration rate

	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*	*-11.018*	*-9.062*	*1.789*
R_2	*0.000*	*-3.903*	*-5.217*	*-11.576*
R_3	*0.000*	*7.071*	*7.354*	*7.463*

**R_1: Estonia, R_2: Latvia, and R_3: Lithuania. See Sections 5 and 6.2 for assumptions. Source: Own calculations.*

Considering estimates in Table 6 it is straightforward to identify that different levels of transport costs lead to consistent estimates of migration flows. While results show substantial differences in the gross migration 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 follow a pattern we assumed in the simulations. As already explained in the previous Subsection, initially there is no migration (column M_{od}^{BR} in Table 6). When regions integrate, Lithuania seems to bee the biggest gainer from trade cost reduction. It steadily gains workers throughout the simulations. 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%.

Latvia seems to be the largest looser from the integration 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%.

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-linear relationship between the level of regional integration and the regions' share of mobile factors. Trade cost reduction among the largest region R_3 and the smallest region R_1 (Lithuania and Estonia) facilitates concentration of manufacturing in the largest region. When trade costs are falling below some critical point (sustain point) a diversified equilibrium becomes sustainable and Estonia - the smallest and peripheral region starts to gain the share of mobile workers (row R_1 in Table 6).

These induced effects – changes in manufacturing price index, wage rate and industrial output level, which work in the reverse direction (from left-hand-side towards the right-hand side variables), are not captured in the standard reduced-type approach. Ability to capture these effects is the main innovation of the 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 migration equation, which is derived entirely from the theoretical NEG model.

Although, the New Economic Geography has a history of more than 15 years, its potential has not been appreciated in the migration literature yet. Using an estimated NEG model we demonstrate the tremendous potential of its structural nature. In particular, because of the endogeneity between right- and left-hand side variables, the predicted migration numbers would be biased in a reduced form model. Moreover, our simulations advocate that even relatively moderate changes in some of the variables (such as transport costs) can actuate unpredictable (both in sign and magnitude) changes in other explanatory variables (such as wages). We are 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 testing structural models in predicting migration flows.

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. However, labour mobility in the Baltic countries is sufficiently low to make the swift emergence of a core-periphery pattern very unlikely at NUTS 2 geographical level. As far as economically motivated migration depends on differences in the level of prosperity between countries (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 population.

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.

References

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

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

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

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.

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

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

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.

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.

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

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

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

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.

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.

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.

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

Hatton T.J. (1995), A Model of UK Emigration, 1870-1913. Review of Economics and Statistics, 77:407-15.

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

Kielyte, J., Kancs, 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.

Krugman, P. (1991), `Increasing Returns and Economic Geography', Journal of Political Economy, 99, 3, 483--99.

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

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

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

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

Pflüger, M. (2004), `A Simple, Analytically Solvable, Chamberlinian Agglomeration Model', Regional Science and Urban Economics (forthcoming).

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

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