Modeling aggregate migration and cross-border worker flows

Ferdy Adam

February 20, 2009 (very preliminary)

Abstract

This paper deals with endogenous foreign labour supply in a small open economy - Luxembourg - which is heavily dependent on it. Foreign labour supply can manifest itself through migrations or cross-border commuting. Three equations are estimated, in the spirit of the modified gravity models: one for immigration, one for emigration and one for cross-border worker flows. Significant r.h.s. variables are, besides labour demand: real net incomes of workers, unemployment rates and house prices. The importance of endogenous foreign labour supply is illustrated through simulations with a large scale macro-econometric model that integrates the three estimated equations. Rendering foreign labour supply endogenous helps to explain the muted impact of demand shocks on unemployment or the relatively recent emigration of residents to the neighbouring regions. It also shows to what extent increases in foreign labour supply can have positive or negative impacts on the domestic economy.

Keywords: Migration, Cross-border workers (Commuting), Macroeconometric modelling

JEL Classification: C50, J61

1 Scope of the article, main findings

As a small country (half a million inhabitants), Luxembourg - to some extent - relies on foreign labour supply in order to grow. In fact, among other reasons, due to the large labour reservoir of the neighbouring regions ("Grande Grégion"), Luxembourg has been able to grow more than twice as fast as the euro-zone area over the last twenty years. The scope of this article is to endogenise foreign labour supply in a standard macro-econometric model describing the Luxembourg economy.

Foreign labour supply can take two forms: (a) immigration, which means foreign workers establishing themselves within the borders of the country, and obtaining a job there; (b) across-the-border commuting which consists in daily travelling to and from Luxembourg. Nowadays, a little more than 40% of all workers having a job on Luxembourg territory are cross-border workers (hereafter: CBW) and more than 60% are non-nationals (non Luxembourgers, see...
The average migration rate has been around 0.7% over the last 20 years. Hence, omitting this endogeneity of migration flows when carrying out forecasting or policy analysis w.r.t. the Luxembourg economy can create an important source of errors or mis-judgments. Simple descriptive statistics reveal the pro-cyclicality of the net migration rate. CBW depend heavily on job creation on the territory. But labour demand or the economic cycle cannot be taken as the sole factors driving labour movements to and from Luxembourg.

In the spirit of the literature on the so-called "modified gravity models" (Greenwood 1997), other factors, such as wages, employment perspectives or living costs (housing, consumption) should play a role. Again, simple descriptive statistical analysis reveals important differences in net, after tax wages between Luxembourg and the neighbouring countries/regions, in housing costs and in unemployment rates. These variables are the most prominent series appearing in modified gravity models in order to represent the pull/push forces that drive migrations.

In that sense, we have estimated three equations explaining in-coming and out-going migrations for Luxembourg as well as CBW movements. They have been estimated as error-correction mechanisms, after some pre-testing (stationarity, cointegration). Data is yearly and a relatively restricted number of observations (28) limits the scope of the findings except for the CBW equation where econometric evidence is of much better quality.

These three equations have been integrated in a large macro-econometric model, describing the Luxembourg economy (Adam 2004, 2007), called Modux. Modux is commonly used at STATEC, the governmental agency for statistics and economic analysis, for forecasting and policy analysis. We argue that the inclusion of these equations in Modux considerably improves the quality of the simulations. To that purpose, we show results for two sets of simulations:

- a strong negative demand shock on the economy, through the decrease in the share of the banking sector in GDP;
- an increase in the differences which prevail between Luxembourg and the bordering regions concerning the main variables governing foreign labour supply, namely net wages, unemployment rates and house prices.

The first set of simulations shows that endogenous migrations generate, in case of a strong negative demand shock, important outflows of labour that considerably dampen the impact on resident unemployment. As a result, when CBW and migration flows are completely endogenous, along the lines of the modified gravity models, unemployment hardly increases as a result of the strong negative demand shock. Only when rendering migrations and CBW exogenous, in the same model setup, resident unemployment does it increase.

The second set of simulations increases foreign labour supply addressed to the Luxembourg economy by emphasizing the already existing differences between local and foreign wages, unemployment and house prices. Higher foreign labour supply then generates a large negative demand shock in the country through
substitution between resident and non resident labour: GDP, wages and prices decrease whereas unemployment increases. In the long run however, the overall impact of these shocks is not completely negative. Indeed, GDP has completely recovered after 15 years, in two of three cases, and the same should be true, with some more delay, for employment.

The remainder of the paper is organised as follows: the second section incorporates a literature survey on the "modified gravity models". The latter are widely used to explain aggregate flows of migrations within or between countries. The third section exposes the theoretical equations that form the backbone for the estimations. In section four, the underlying data is described while the fifth section contains the results of the estimations. Finally, section six reveals the results of the simulations while the last section concludes. The appendix contains a.o. a more detailed survey on the literature concerning the modified gravity models.

2 Literature review

2.1 Definitional aspects

The goal of this study is to explain the flow of aggregate migrations to and from Luxembourg. These migrations include traditional moves involving a change of residence and of workplace as well as cross-border worker movements i.e. journeying from place of residence to workplace, involving more or less considerable distances. Both types of movements constitute foreign labour supply addressed to the Luxembourg economy.

Firstly, some definitional aspects will be clarified. Indeed, in the case of a small regional economy like Luxembourg, a "migration", i.e. the residential move of a French worker to Luxembourg - crossing a jurisdictional border (covering for example 20 km) - might not be driven by the same theoretical model (or economic variables) as that of an Indian migrant to the UK.

The common literature retains at least four notions of traditional migrations:
- the "standard" international migration, i.e. the more or less permanent move involving the crossing of mostly large distances, often in the context of high income per capita discrepancies between the originating (poor) country and the (rich) host country;
- what Cushing and Poot (2004) call "temporary circulation", i.e. migration of mostly highly qualified professionals, driven by globalization, which leads to an intensifying exchange of workers, for relatively short periods of time (several years);
- internal or regional migration within single countries; the latter being necessarily large, with the US as an example, for which an important amount of literature exists;
- residential moves, which are defined as changes of places of residence without changes of workplace.

As a matter of fact, all four notions apply to Luxembourg (but more so the
last three) and can hardly be disentangled within the data. On the other hand, these definitions do not directly take into consideration cross-border movements or commuting. In order to define commuting (or to distinguish residential moves from true migrations - i.e. involving a change in the place of work) it is crucial to define functional labour market areas (LMA, as in Cöppers and Hensen, 2003, for example). These authors define functional regional labour markets by modeling commuter behaviour in the NL. They use statistical criteria to subdivide the regional labour markets, overcoming administratively defined areas that are often used as units of analysis. Indeed, labour market policies targeted at administratively defined areas may be less effective if the administrative boundaries do not follow functionally linked labour market areas.

Building upon the notion of functional labour market areas, Eliasson, Lindgren and Westerlund (2003) suggest an explicit definition of commuting or migration: according to them, geographical labour mobility involves transitions in at least 3 dimensions: (a) residential change, (b) change of workplace, and (c) transitions between different labour market states (unemployed, employed, etc.). On the other hand, geographical mobility is based on labour market areas (LMA) which are functional regional units where most people can find both a place to live and a place to work. Mobility between LMAs tends to be motivated by labour market reasons rather than housing market considerations. Hence interregional labour mobility (between t and t+1) can be defined in the following ways:

1) for those that had a job in t:
   a) change of workplace and of residence;
   b) no change of residence and change of workplace by starting to commute;
   c) no change of region of residence and change of commuting destination;

2) for those that did not have a job in t, found a job in a workplace in t+1 different from their LMA in t, and:
   d) migrate, or
   e) start commuting.

Hence: Migration = (a) and (d); commuting = (b), (c) and (e).

In this study, we try to use the above mentioned definitions in order to post some clarifications. Firstly, one can assume that there is only one functional LMA, namely the country Luxembourg plus the bordering regions from where most commuters are drawn. Commuters or cross-border workers are defined according to the fact that they cross a jurisdictional border, not according to the fact that they commute between LMAs\(^1\). Residents that move to another country (in order to enjoy cheaper accommodation) but that stay in the same LMA hence become commuters. On the other hand, residents of neighbouring countries might move closer to Luxembourg (but continue to reside in their home country) in order to take a job there: first they undertake a residential move within their country of origin to become commuters. If they decide to

\(^1\)Although some might: the neighbouring cities of Metz, Thionville, Trier and Saarbrücken probably consist of true LMAs.
move still closer to their place of work, and, by doing so, cross the border to become residents of Luxembourg, they generate a migration, which, in the sense of the literature commented above, consists of a residential move (they stay in the same LMA).

Recently, an increasing number of residents of Luxembourg is moving to the neighbouring countries, mostly in order to find cheaper accommodation. These workers still stay in the same LMA, their move is a residential one but they become cross-border workers. According to Pigeron (2008), in 2005, 2100 inhabitants, that already lived in the country in 1995, had moved to the bordering regions. If this number seems small, Pigeron observes that, while flows in both directions exist, those of CBW moving into the country were decreasing while emigration flows were increasing. According to other sources\textsuperscript{2}, in 2006, 7050 Luxembourgers (of nationality) where living in neighbouring german and belgian villages. While the number of Luxembourgers living in Belgium was quite stable between 1995 and 2006, those living in german municipalities has been multiplied by 2.2.

There are some papers that discuss the joint (individual) decision whether to migrate or to commute. Eliasson et al. (2003) provide a structured view on the differences between migration and commuting by examining the determinants of the individual’s decision whether to search for and accept a job requiring interregional mobility and the interrelated choice between migration and commuting as the mobility mode. So et. al. (2004) apply a restricted multinomial logit framework to the migration/commuting decision in and around Des Moines (Iowa/USA) by designing an empirical model of joint decisions of where to live and where to work: they find that individuals make residential and job location choices by trading off wages, housing prices and commuting costs. Romani et al. (2003) estimate a simultaneous, discrete choice model of commuting and moving; they analyse the effects of individual characteristics of workers on their commuting and migration decisions by drawing comparisons between Spain and other countries.

However, due to the nature of our data (gross migration flows, on a country-by-country basis, to and from Luxembourg), the literature that is of interest consists of those papers that explain gross (not individual) migration (and commuting) in a regional context, be it international migrations (crossing administrative borders) in areas such as the European Union with small countries and greater likelihood of migration (involving crossing of national borders) or be it regional migrations in larger countries such as Canada or the United States. Cushing/Poot (2004) argue that models of gross interregional migration can be equally fruitfully applied to international migration where international migration flows are fairly unrestricted such as within the European Union or Australasia. It is a matter of fact that the prevailing part of migrants into Luxembourg comes from EU27 countries, in the context of virtually unlimited labour flows. In that sense, the literature that is of most interest for us is the one taking as

\textsuperscript{2}Luxemburger Wort, 16.5.2007, citing a "Question to Parliament" and the respective answer by the Minister of Economics, which is based on population registers in neighbouring countries.
a starting point the modified gravity model or some similar model.

2.2 Gravity models

According to Maré and Choy (2001), the simple gravity model applied to migration streams is based on the Newtonian law of gravitation: the force of attraction between two bodies (or migrations $M_{ij}$ between two regions/countries) is proportional to the product of the masses (population $P_i$ and $P_j$) and inversely proportional to the squared distance $D_{ij}$:

$$M_{ij} = G * P_i * P_j * D_{ij}^2$$

$G$ being a constant.

The model can easily be generalized (cf. Greenwood 1997) by expressing (1) in double-logarithmic form and abandoning the fixed elasticities on the mass (population) and distance variables:

$$M_{ij} = G * P_i^\alpha * P_j^\beta * D_{ij}^\delta$$

(2) is seldom applied in this form. Gravity models are most of the time augmented (and called modified gravity models, MGM) through the inclusion of variables that have a behavioural or economic content: income/wages, unemployment rates, house prices, weather/climate, degree of urbanization, heating costs, measures of public amenities, etc... According to Greenwood (1997), MGM hold an important place in the migration literature because their formulaters tried to incorporate behavioural content in the context of the gravity model approach.

An important drawback of the MGM is the fact that they are not based on proper theoretical foundations. Greenwood (1997) and Maré/Choy (2001) advance some other criticisms. There is however a certain amount of published papers that apply this type of model to estimate aggregate migration flows: Foot and Milne (1984), Pissarides and McMaster (1990), Daveri and Faini (1999), Gorbej, James and Poot (1999), Karemera, Oguledo and Davis (2000), Parikh and Van Leuvensteijn (2003), Mathä and Wintr (2007), Marvakov and Mathä (2007) and Ashby (2007). These papers are reviewed in the appendix; a short summary concerning the respective theoretical background for the use of the MGM as well as a more detailed explanation of the modified gravity model are given hereafter.

Following Greenwood (1997), the modified gravity model is an augmented version of (2): the basic gravity model, derived from physics, is improved through the inclusion of economic variables $X_i$ and $X_j$:

$$M_{ij} = G * P_i^\alpha * P_j^\beta * D_{ij}^\delta * X_i^\lambda * X_j^\mu$$

The rationale for the inclusion of economic or behavioural variables in the MGM is diverging: In Foot and Milne (1984), Mathä and Wintr (2007) and Marvakov and Mathä (2007) there is no real theoretical point of departure at all:
the MGM is simply applied to the data. Pissarides and McMaster (1990) argue that household utility is based on costs and gains of moving (migrating) the latter being assumed to depend on wages and unemployment: "in a time series migration regression, relative wages and unemployment differentials are likely to be the most important explanatory variables since these are the determinants of migration that show most variation over time". Daveri and Faini (1999) build on the portfolio approach to the determination of the optimal family size as developed by Appelbaum and Katz (1991). Karemera, Oguledo and Davis (2000) rely on Borjas (1989) that assumed that a decision to migrate depended on expected payoffs from migration as well as on migration costs. According to Karemera, Oguledo and Davis (2000), Borjas (1989) summarized his theoretical model of immigration in three equations but failed to provide estimated values of the suggested elasticities. They therefore simply formulated and estimated a gravity model explaining international migration. The see the MGM as a reduced form equation derived from a system of demand and supply relationships. In Parikh and Van Leuvensteijn (2003) the theoretical foundation is the "option value approach" to migration as put forward in Dixit and Pindyck (1994), Burda (1993) and Burda et al. (1998). According to those, the opportunity cost of migrating today, in addition to the expected net present value of future income gains from migration net of migration costs, is referred to as the option value of waiting. Finally, Ashby (2007) applies spatial econometrics to a gravity model by using a contiguity matrix that measures proximity to a given observation/state.

The economic variables that are mostly used in these MGM are income (or proxies thereof: 8 occurrences) and unemployment (8). Other variables (other than population or related and distance) included in MGM are employment (size, growth: 3), language (3), inflation (2), human capital endowment (2) and a measure of risk (1).

### 3 Theoretical model

Most of the models estimated in the above-listed reference papers consist of some sort of double-log equation where migration is on the l.h.s. Hence, in this paper, (3) is estimated in a simple double log-linear form. Assuming that there are only two functional labour market areas (Luxembourg and the neighbouring regions, taken as one single region), the distance term \( D_{ij} \) can be skipped (because constant). The population variable of the sending countries displays much less variance than the one from the receiving country (Luxembourg). It can be reasonably assumed that it is constant and is skipped therefore (becomes part of the constant).

Due to a reduced set of observations (1980-2007), it seems preferable to apply the same elasticity to the same domestic and foreign economic variable \( X (\lambda = \mu) \). (3) becomes then:
\[
\log(M_{ij}) = c + \beta \log(P_j) + \lambda_1 \cdot \frac{X_{1,i}}{X_{1,i}} + \cdots + \lambda_n \cdot \frac{X_{n,j}}{X_{n,i}} + \varepsilon \quad (4)
\]

where \(i\) stands for the sending and \(j\) for the receiving country (Luxembourg).

Economic variables \(X_1\) to \(X_n\) to be included in the final model are employment (labour demand), labour income, unemployment and house prices. The rationale for including labour income and unemployment seems straightforward.

House prices (or related variables) are integrated in a few papers dealing with migrations\(^3\). The idea is to take into account high existing differences between Luxembourg and the neighbouring countries: substantially higher house prices in Luxembourg than in the neighbouring regions/countries could generate more cross-border movements w.r.t. classical migrations (i.e. settling in the country while taking a job there).

If migrations respond to economic conditions, labour demand should be an interesting candidate variable, although in the literature, it has not been used as often as income and unemployment. In our case, the argument is the same as with (foreign) population: foreign employment shows much less variance than domestic employment and we stick to using domestic employment only.

For homogeneity reasons, it seems preferable to fix \(\beta\) (the elasticity on domestic population) to one - resulting in having the migration rate as dependent variable - and to take the ratio of domestic employment with respect to total population as some sort of long-run anchor for the migration rate, w.r.t. the evolution of domestic employment. In the case of the cross-border worker equation, \(\phi\) (the elasticity w.r.t. total employment or labour demand) has been set to one, again for homogeneity reasons, and domestic population can then be dropped.

Finally, three models have been estimated for Luxembourg, in agreement with the MGM theory:

- one explaining immigration;
- one explaining emigration;
- one explaining cross-border worker movements:

\[
\log\left(\frac{M^k}{P_j}\right) = c + \phi \cdot \log\left(\frac{L_j}{P_j}\right) + \lambda_1 \cdot \frac{X_{1,i}}{X_{1,i}} + \cdots + \lambda_n \cdot \frac{X_{n,i}}{X_{n,i}} + \varepsilon \quad (5)
\]

where

- \(M^k\) are migration movements (formerly \(M_{ij}\)), where \(k=m,e,c\) standing for either immigration (from foreign country Luxembourg), emigration (from Luxembourg to foreign country) or cross-border workers (foreign to Luxembourg);
- \(P_j\) is domestic population;
- \(L_j\) domestic employment, and
- \(X_1\) to \(X_n\) independent economic variables (income, unemployment and house prices).

\(^3\)Eliasson et al. (2003), Cannari et al. (2000).
4 Some comments on the data used in the regressions

All series used for the regressions are annual series. Estimations start in 1980 but some series go back to 1970. Data is in general "by region", but the definition for "region" is not homogenous: it is not the same for example for house prices than for unemployment. Data on net wages is "by country" and has been deflated in the regressions by relevant consumer prices.

4.1 Dependent variables

There are three series which are treated as dependent variables, i.e. that figure on the r.h.s. of the estimated equations: the immigration and the emigration rate (fig. 2) and the ratio of cross-border workers to domestic employment (fig. 6). These series are all produced and/or published by STATEC. Whereas migrations, in or out, are flows, the series for cross-border workers is a stock, namely the total number of non-resident workers having a job in Luxembourg. International civil servants as those employed by the European Commission are excluded from domestic (interior) employment. Migration data as we use it exists from 1980 onwards, the crossborder series starts in 1970.

4.2 Independent variables

There are four independent series that are used in the regressions: domestic employment or labour demand, net wages, unemployment and house prices. In the regressions, the latter three are used as ratios, i.e. the home series over the foreign series (fig. 14).

Net wages, for the three neighbouring countries B, F and D, are obtained from the OECD publication "Taxing wages": data could be traced back until 1972 (fig. 9). The net wage (i.e. after taxes and social transfers) is for a typical married worker, with two children and corresponds to 100% of the average wage.

It should be obvious that worker-consumers react to real wages or purchasing-power: hence, net wages have been deflated by respective prices indices: the domestic or Luxembourg consumer price index for the wage earned and spent in Luxembourg and the foreign consumer price index, which is proxied by the euro-zone GDP deflator, for the neighbouring regions. One more precision: CBW spend part of their income in Luxembourg (around 25%\textsuperscript{4}). Hence, the net wage earned in Luxembourg and entering the CBW regression has been deflated by a weighted average of the Luxembourg consumer price index and the euro-zone GDP deflator, where the respective weights are 1/4-3/4.

\textsuperscript{4}Cf. Langers and Schuller (2005) and Zanardelli (2005). According to figures drawn from both contributions, it can be inferred that the average CBW spends about 25% of it's net income in Luxembourg (gross annual wage in 2003: 41500 EUR; presumed average net wage, after taxes and social contributions: 31100 eur; average annual spending in Luxembourg: 8000 EUR).
Unemployment data is from Eurostat/New Cronos and comes from the regional exploitation of the national labour force surveys. The series start in 1983 and the regional split-down is chosen so as to capture the areas where most cross-border workers live (cf. fig. 7).

House price data has different national sources. Data for Luxembourg and Belgium is taken from the deeds executed by notaries as transmitted to fiscal and/or statistical authorities. Whereas data for Belgium contains information on the size of the objects (square meters), data for Luxembourg does not. Hence, for all countries, prices are "per object" (i.e. appartments or houses) and not per surface. Data for Belgium concerns the region of Wallonia and the series for Luxembourg and Belgium start in 1980 but the last observation (2007) is missing for Luxembourg (cf. fig. 12). Hence a general point: where series for house prices are missing, data has been back- or forward-spliced with prices for construction land (cf. fig. 11).

House price data for France concerns only new buildings as brought to the market by promoters and sold to private housholds. It excludes transactions on existing, new or old, buildings. Data on these prices goes back only until 1994\(^5\). There is not more information on prices for construction land: hence, series for other regions have been used to obtain data points before 1994.

Apparently, the only source for regional house prices for Germany are by city price-brackets, as collected and published by professionals \(^6\). This data goes back to 1990 for Saarland and 1995 for Rheinland-Pfalz. Prior data points are obtained with prices for construction land.

In general, despite the different sources for house prices (besides, of presumably rather poor quality), both levels and variations seem congruent with other facts and economic evidence:
- prices are notably higher in Luxembourg; there was however no significant difference yet at the beginning of the 90s;
- prices in the german regions have been stagnating for more than ten years, or even falling, a tendency which has to be seen in relation with the stagnating german economy (until 2005) and the depressed housing market;
- prices have somewhat accelerated in the bordering french and belgian regions, although our split-down is too broad to single out the price-boom which happened in areas much closer to the Luxembourg border.

\section{5 Estimation results}

Before estimating error-correction equations and/or carrying out cointegration analysis, series in (5) should be tested for their order of integration. It has been found that all series except one entering (5) seem to be I(1): detailed test results

\(^5\)Enquête sur la commercialisation des logements neufs*

\(^6\)For example, concerning the Saarland: "Immobilienpreisspiegel Saarland" as published by the respective Immobilienverband, which is the professional organisation grouping real estate agents.
are in the appendix (table xy). Only for the ratio of cross-border workers to total employment, evidence that the series is I(1) rather than I(0) is weak.

As a second step, the level equations have been estimated by OLS. MK (2002) argue that in general, in order to derive the long-run parameters, static (i.e. level only) regressions are to be avoided (cf. op. cit., p. 183-84). A preferred method would be the estimation of the complete dynamic equation. This has been done - and it works well - for the CBW equation (cf. hereafter), but not for migrations. Hence the static estimation of the long-run parameters, which is probably the second-best option, has been undertaken in first place. It comes down to the Engle-Granger two-step procedure (as a means of testing cointegration), although a specific method is applied (cf. Zivot 2000), namely the one-step estimation of the "dynamic" system, with the long run parameters fit in the difference equation (instead of the residuals from the static regression) and the cointegration test as a standard t-test on the lagged long-run parameter of the dependent variable (cf. MK, p. 205). Zivot argues that "the ECM t-test with a pre-specified cointegrating vector can have much higher power than single equation tests for cointegration based on estimating the cointegration vector".

Estimation results for the static, long-run equations are in table xy. An ADF test for the stationnarity of the residuals, with the McKinnon critical values, as cited by Maddala and Kim (op. cit., p. 201), reveals that neither of the estimated equations denotes a cointegrated relation.

The immigration ratio (immigration over total population) has been estimated with and without constraining the parameter on employment (total employment over total population) to 1. In both specifications, the ratios of unemployment rates and net real wages have the correct sign: more unemployment means less immigration and higher wages attract more immigrants. House prices show up with a (significant) positive coefficient but the estimated values for the other independent variables change sign and/or take counter-intuitive values when house prices are integrated in the regression. Two problems might exist: high collinearity between house prices and the other independent variables and/or an endogeneity bias: simple econometric and statistic evidence seems to confirm that house prices depend on the immigration rate resp. domestic revenues. Due to these reflections, relative house prices have been taken out of the immigration equation.

House-prices show up significantly in the CBW equation and have also been retained for the emigration equation. For the latter, unemployment seems less a determinant: once foreign workers are "in", they seem less deterred by a worsening of the situation on the labour market. House prices have a positive sign, meaning that higher house prices are driving residents out of the country.

Concerning the CBW equation, the estimation of the level equation has less interest since the estimation of the dynamic ECM works well. Although
values for the parameters of the independent variables differ quite largely between specifications (level only, dynamic, constrained or unconstrained), the more qualitative evidence remains stable: all variables have generally the expected sign and are higher in absolute value than for the migration series. CBW movements seem more reactive to the economic situation in Luxembourg, something quite intuitive since it is easier to change the work-place than the country of residence. This being said, it should be kept in mind that CBW is a stock variable whereas migrations are flows, i.e. differences of stocks: it seems hence quite intuitive that absolute values of coefficients on flow variables are lower than those on stock variables.

In a second step, all equations have been estimated in one step: results are in table xy. Concerning the two migration equations, the long-run parameters of the static OLS regression have been fitted into the ECM. This procedure serves as an implicit cointegration test and is described in Zivot (2000) and in Maddala and Kim (2002). The test statistic is on the long-run, lagged, dependent variable: a standard t-test suffices as a check against cointegration. Thus, test statistics, as reported on the error correction term in table xy are different for cross-border workers than for migrations. Furthermore, estimation results seem to indicate that the immigration equation is based on a true cointegration relation whereas emigration is not. Concerning CBW, the dynamic, one-step estimation of the ECM and the application of the Banerjee (1998) test statistics clearly indicate a cointegrated relation.

Some further comments regarding the long-run elasticity of CBW w.r.t. total employment. Static OLS seem to indicate that the fixation of this parameter to one results in a more significant test-statistic concerning the residual based test for cointegration (but the difference is very weak and the test-statistic is any way below the threshold). On the other hand, rolling regressions indicate that this parameter is converging towards one: indeed, as the share of CBW in total employment is increasing, the elasticity (of CBW w.r.t. total employment) necessarily approaches one. Hence, in the dynamic estimation, this parameter has been fixed to one. In general, the CBW equation seems quite stable, as out-of-sample forecasts indicate.

Summing up, it can be argued that the econometric derivation of the parameters for CBW movements and migrations was a mixed success: results are quite satisfactory for CBW but less so for migrations, although the values of the estimated coefficients are quite meaningful. Further work should primarily be related to the endogenous treatment of house prices. Simulation analysis in the next part will constitute a further test on the plausibility of the obtained relations.

6 Simulations

The goal of this part is to illustrate the importance of endogenous foreign labour supply through simulations (and implicitly to test the plausibility of the estimated relations).
To that purpose, the three equations are embedded in a larger, estimated macro-model describing the Luxembourg economy, called Modux (Adam 2004 and 2007). Modux is regularly used for forecasting and policy simulation purposes within STATEC. Modux is a fairly standard macro-econometric model, with more than fifty behavioural variables.

Instead of simulating the three labour supply equations separately, they are embedded in the model which is then simulated as a whole, in order to reproduce all feedback mechanisms. Two sets of shocks are produced: the first one substantially decreases the part of the banking sector in the Luxembourg economy, with respect to a given baseline. The baseline is a "neutral" projection of the Luxembourg economy until 2020\(^8\). The second one explicitly shocks the three exogenous variables that influence foreign labour supply, namely foreign unemployment, foreign net labour income and foreign house prices. All three variables are modified in a sense so as to increase existing differentials with Luxembourg: wages and house prices are decreased whereas unemployment is increased.

The results of the first simulation can be seen on figures 15 to 23 in the appendix. Figure 15 shows the baseline and the deviation of it for the share of banking sector value added in GDP. In the simulation, the latter is kept stable around its 2010 value, namely some 23\% (of GDP). In the baseline, it rises to about 29\%, which comes close to an extrapolation of the rising linear trend observed since 1980.

In order to gauge the impact of endogenous foreign labour supply, the simulations are also carried out with "exogenous" migrations and CBW. In fact, whereas migrations, in and out, are truly kept exogenous (hence population growths along a fixed path, whatever the outcome of the economy), CBW become a fixed share of total dependent employment, namely the value of the share observed in 2007, some 43\%. Hence, there is no influence on CBW from the three economic variables that govern foreign labour supply in the normal, standard case.

As a consequence of the strong decrease in banking sector activity, GDP, in volume terms, loses around 6-7\%. This corresponds to a decrease in the average growth rate of about 0.5 percentage points per year. The most striking result is the fact that unemployment increases only when migrations are exogenous: it rises some 0.6 percentage points at most. This is due to the huge outflows of labour when migrations are endogenous: the decrease in (real) domestic wages, the (initial) increase in unemployment and the fall in labour demand trigger outflows of workers, both residents and CBW.

Outflows of labour are so important that resident unemployment decreases after 9 years, a result which is somewhat counter-intuitive. A detailed analysis shows that this decrease is due to the fact that total and working-age population

\(^8\) Basically, it is an extension of the official forecast which stops in 2010. Stationary exogenous variables are extended at their latest observed value whereas non-stationary variables recover their past, long-run trend growth rate. Some corrections had to be made to stabilise the results (for example, the participation rate had to be topped at a certain level, despite not being stationary, i.e. trending upwards since the beginning of the eighties).
experience, after 2015, a stronger decrease than resident employment, which stabilises (figure 18). The stabilisation of resident employment comes from the stabilisation of total employment, respectively of GDP, whereas the continued decrease in total population is linked to the persistent outflow of labour through migrations.

One might however expect that this decrease of resident unemployment is only temporary, as the deviation of the total migration rate from baseline is starting to become smaller after 2015 (figure 19). An important drawback of Modux is indeed its inability to produce true long-run results, in the sense that deviations from baseline of the growth rates of all variables do not stabilise around 2020 and that one should be able to extend the simulation range in order to see a clearer picture.

What can be the conclusions from this first set of simulations? Endogenous migrations produce more realistic outcomes in the sense that a strong negative shock on output and employment is likely to see outflows of labour. As a result, the impact on unemployment is diminished. The reverse seems also quite intuitive: as a result of a (strong) positive demand shock, foreign labour increasingly flows into the country and mitigates the effects on domestic employment and unemployment.

The second set of simulations sees a modification of the three exogenous variables that govern foreign labour supply: revenues, unemployment and house prices.

In the short run, in all three cases, there is a decrease in resident employment and an increase in resident unemployment (figures 27 and 28). This is mostly due to an increase in net migrations and CBW, except for the case where foreign house prices are decreased: the latter induces a reduction in net migrations but a strong increase in CBW. The "reduction-in-foreign-house-prices" shock is explained more in detail somewhat later.

The decrease in resident employment and increase in resident unemployment, together with the increase in foreign labour inflow, can be interpreted as a substitution of resident workers for non-residents. This substitution mechanism is however not explicitly modelled in Modux, since there is no distinction, in terms of factors of production, of resident and non-resident workers (CBW, migrants). The consequences of the decrease in resident employment and the increase in resident unemployment are lower private consumption, GDP and wages.

In the medium term, some impacts on some variables change sign. Firstly, the shock on foreign unemployment decreases net migrations (formerly, they increased) whereas the impact on CBW becomes stronger. This is due to the fact that CBW are more sensitive, in the long run, to unemployment differentials than migrations. The same reason holds in order to explain the decrease (instead of an increase in the short run) of CBW when foreign net revenues are shocked: whereas this should in principle increase foreign labour supply, namely CBW and incoming migrations, the increase in resident unemployment, at some point, reduces CBW, w.r.t. the baseline, and the increased supply of foreign labour exclusively comes from migrations. As a result, since increased migrations mean
increased population, resident employment increases after a while.

Still the overall impact on the whole economy remains negative, but GDP slowly revovers to the baseline value, in all three cases, with different speed though, despite the increase in resident unemployment which does not vanish.

The decrease in foreign house prices is an interesting case. It describes a phenomenon which has manifested itself in the Luxembourg economy since the beginning of this decade, namely the outflow of residents, foreigners or nationals, to the bordering regions, in order to take advantage of lower housing costs. While keeping their job in the country, they become CBW: according to Pigeron (2008), in 2005, around 0.5% of the inhabitants that had a job and lived in the country in 1995 had left to the bordering regions. This represents a much higher fraction of resident employment (around 1.25%).

In the three estimated equations, house price differentials manage to explain this phenomenon: they appear with a positive sign in the emigration equation and in the CBW equation, with a much higher value in the latter. Hence, an increase in house-price differentials (which is the result of lower foreign prices, cf. the shock, but which corresponds also to what has been happening since the beginning of the eighties) triggers a decrease in net migrations and an increase in CBW. To explain the outflow of several thousand residents since 2000, simulations show that if house prices in Luxembourg had grown at the same pace than those in the neighbouring regions (i.e. at 3.2% per year between 2000 and 2007 instead of 9.8%), all other things being equal, domestic population would be higher in 2007 by around 9000 persons.

Some tentative conclusions from the second set of simulations. According to Pierrard (2008), "under certain conditions, increasing competition from CBW may generate positive externalities on vacancy openings and reduce domestic unemployment". Our simulations seem to refute this result, since resident unemployment rises in all three cases, as a response to increased foreign labour supply. On the other hand, the results of Pierrard (2008) have been obtained with a DSGE model, and are therefore true long-run results. The simulations carried out with Modux however cannot be extended far enough (i.e. beyond 2020) in order to produce true long run results, and therefore comparisons between both cases seem delicate. Still, the simulations with Modux do show some positive effects on the domestic economy, namely, in all three cases, an increase in GDP (w.r.t. baseline), after some 15 years, as well as an increase in total employment, in two cases. It would still be interesting to test whether simulating Modux over a longer horizon would result in a decrease in resident unemployment, as in Pierrard (2008).

On the other hand, Modux, through it's error correction mechanisms, seems better suited to produce short-run results, which, regarding policy makers, are equally - if not more - interesting than real long-run, steady-state results. In that sense, a policy which increases house-price differentials or net wage differentials...
with the neighbouring regions, and which supposedly, at least in the short run, increases resident unemployment, should be considered carefully.

7 Conclusions

The scope of this article was to endogenise foreign labour supply in a standard, large, estimated macro-econometric model describing the Luxembourg economy. Foreign labour supply can take two forms: (a) immigration, which means foreign workers establishing themselves within the borders of the country, and obtaining a job there; (b) across-the-border commuting (or cross-border workers, CBW) which consists in daily travelling to and from Luxembourg. We have argued that omitting this endogeneity of migration flows when carrying out forecasting or policy analysis w.r.t. the Luxembourg economy can open up an important source of errors and mis-judgments.

In the spirit of the literature on the so-called "modified gravity models", we have estimated three equations explaining in-coming and out-going migrations for Luxembourg as well as CBW movements. Independent variables are net real wages, employment (labour demand), unemployment and house prices. These variables are the most prominent series appearing in the literature on modified gravity models in order to represent the pull/push forces that drive migrations but modelling aggregate CBW flows along the lines of these models is quite novel.

These equations have been estimated as error-correction mechanisms and have been integrated into a large macro-econometric model, describing the Luxembourg economy. Econometric results are quite satisfactory for the CBW equation but much poorer for the migration equations.

We then show that the inclusion of these equations in Modux considerably improves the quality of the simulations. To that purpose, we analyse two sets of simulations:

- a strong negative demand shock on the economy, through the decrease in the share of the banking sector in GDP;
- an increase in the differences which prevail between Luxembourg and the bordering regions concerning the main variables governing foreign labour supply, namely net wages, unemployment rates and house prices, besides labour demand.

The first set of simulations shows that endogenous migrations generate, in case of a strong negative demand shock, important outflows of labour that considerably dampen the impact on resident unemployment. As a result, when CBW and migration flows are completely endogenous along the lines of the modified gravity models, unemployment hardly increases as a result of the strong negative demand shock.

The second set of simulations increases foreign labour supply addressed to the Luxembourg economy by emphasizing the existing differences between local and foreign wages, unemployment and house prices (wages and house prices are
higher in Luxembourg whereas unemployment is lower, hence a considerable force to attract foreign workers). Higher foreign labour supply generates a large negative demand shock in the country through a substitution effect between resident and non resident labour: GDP, wages and prices decrease whereas unemployment increases. After 15 years, GDP has almost completely recovered and one can presume that the same will be true for employment, with some delay. In general, the simulation horizon seems not long enough to produce true long-run steady state results as in Pierrard (2008). That could be a point for improvement. Other issues that would deserve better treatment are:

- the endogenization of house prices: the latter presumably depend on migration or demographic pressure and their inclusion, as an independent variable, is not without concern (hence they have not been fitted into the immigration equation); some further work would be to estimate a system of simultaneous equations, including a regression explaining house prices;

- the statistical series for house prices could equally be improved, since it does not incorporate information on the quality (size) of the objects;

- whereas the CBW equation could be estimated quite successfully, the econometric results for the migration equations are somewhat poor, since no cointegrating relation could be found;

- since the migration equations could not be estimated successfully, one idea would be to estimate a true gravity model for migrations, along the line of Mathä and Wintr (2007) would did exactly this, but for CBW;

- higher foreign labour supply probably displaces, at least to some extent, resident workers: there hence must be some substitution effect; to analyse this question more thoroughly, CBW could be treated as a specific factor of production.

8 References


9 Appendix

9.1 Equilibrium vs disequilibrium approach

According to the disequilibrium approach, which is historically the first to have been developed, migrations are assumed to be driven (a. o.) by the existence of a set of non market-clearing regional wages. Spatial variations in earnings are supposed to reflect opportunities for utility gains. On the other hand, the proponents of the equilibrium approach assume that spatial variations in wages or earnings are compensating (for other differing variables) and therefore do not reflect opportunities for utility gains.

Underlying the disequilibrium approach is the simple income-leisure model of labor economics where the consumer maximizes utility, subject to an income constraint. The individual is expected to offer his services in the market with
the highest wage which might require migration. The human capital approach (Sjaastad 1962) added to the disequilibrium approach. The potential migrant would select that locality at which the real value of the expected net benefit (i.e. the discounted value of the expected future stream of pecuniary returns), that accrued to him from migration, is greatest.

The proponents of the equilibrium approach assume that households and firms are in equilibrium at any point in time. Utility is made spatially invariant by migration. Any location offering extra-normal utility will experience in-migration until, in some combination, wages fall or rents rise sufficiently to eliminate the utility differential and migrations. Amenity-rich regions (climate, geography, public goods) must have, in equilibrium, some combination of lower wages and higher rents. Net in-migration to amenity-rich areas tends to drive down wages and push up the prices of locally produced goods and services (i.e. rents). Wages and local prices diverge across regions until they just compensate households for the differing amenity bundles.

For equilibrium to prevail, markets must be efficient so that wages and prices quickly realign to clear markets subsequent to any disequilibrating exogenous disturbance. Thus, a reasonable question relates to the speed at which equilibrium establishes in regional goods and factor markets. In the case of Luxembourg, the existence of important migration streams, both in and out, including commuters, seems to refute the equilibrium approach. But the question of the speed at which regional market equilibrium is reached is a very important one. This paper could provide some insights as to how quickly convergence in factor or goods and services markets is reached within the context of the "Grande Région". According to Greenwood (1997), the equilibrium theorists believe that the adjustment should be relatively quick but empirical studies (Greenwood 1991, Pissarides and McMaster 1990, cited by Greenwood 1997) do not suggest extremely rapid adjustments.

9.2 Modified gravity models: detailed review

Foot and Milne (1984) employ a multiregional framework in which net migration over time to all 10 Canadian provinces is analyzed within an integrated system of equations. An extended gravity model is the basis for the equation specification and the use of constrained econometric estimation techniques allows for the provincial interdependence of the migration decision. Any person exiting from one region must enter some other region and hence the sum of all out-migrants must equal the sum of all immigrants. The implications of this requirement are the introduction of cross-equation constraints on the parameters of the estimated regional equations. Bi-regional analysis ignores this system wide constraint (excluding some or all of the independent variables relative to the other regions) and may hence result in incorrect parameter estimates. Any comprehensive analysis must be carried out in a multiregional framework.

Data is based on family allowance payment statistics for Canada for the 1960’s and 1970’s, concerning all 10 regions/provinces. Time-series of origin-destination matrices are available both annually and quarterly. The basic source
is the quinquennial census; annual intercensal and postcensal estimates by province are published by Statistics Canada.

Foot and Milne (1984) estimate an extended, multi-regional gravity model which specifies that net migration in any province \( (NM_i) \) is determined by the size of the population, the real wage and the unemployment rate in that province and distance weighted aggregates of these variables in the remaining nine provinces.

\[
NM_i = f_i(X_i, XO_i)
\]

where \( X_i = [\text{POP}_i, \text{WR}_i, \text{UR}_i] \) and \( XO_i = [\text{POP}_0, \text{WRO}_i, \text{URO}_i] \)

and the \( X_i \)'s are population, wage rate and the unemployment rate of the province that has in-migration and the \( XO_i \)'s are the corresponding (distance weighted) variables for all the other provinces.

The estimated coefficients on the own population variable are all negative and significantly different from zero. Three quarters of the estimated coefficients on the real wage rate variables are found to be of the correct sign and statistically significant. Concerning the unemployment rate, 16 out of 20 of the estimated coefficients are of the correct sign. The results suggest that the extended multiregional gravity model outlined in this paper provides a useful multiregional framework in which to analyze net interregional migration in Canada and elsewhere. With regard to the year of publication of this article (i.e. 1984), it is needless to say that in 1984, matters related to degree of integration and co-integration where (understandably) not treated by the authors.

Pissarides and McMaster (1990) implicitly test the dominance of either the equilibrium or the disequilibrium approach for explaining migrations: regions with above-average unemployment should, in the long run, have above average wages, and vice-versa. Their analysis is built around three questions of which only the first one is of relevance for this research:

- to what extent do regional migration rates respond to regional wage and unemployment differentials?
- to what extent do regional wage and unemployment differentials compensate for each other, in the face of such migration?
- what are the implications of these adjustments for regional economic policy?

The theoretical starting point is the derivation of the net migration rate for a region by considering the determinants of the probability of migration of a single household. They rely mainly on the work by Greenwood as it has been summed up in Greenwood (1985) or Greenwood (1997), the latter being posterior to their article. The net migration rate is a rising function of the gain from moving into the region which depends on relative wages and regional unemployment rates.

The data is for Great-Britain, for the years 1961-1982, and it concerns nine regions which are not true [functional] labour markets but administrative groupings. They use net migration of people of all ages as depending variable. Independent variables are regional wage differentials (i.e. the relative difference of hourly earnings for manual male workers in each region with respect to the national average) and unemployment differentials, which are measured as ratios.
They estimate two equations: one having the net migration rate as dependent variable and another one having regional wage differences as dependent. They pool the data (since they have only 20 observations) by running one single regression for all regions with different constants for each region (thus taking into account region-specific fixed effects). The specification of the equation was selected after a general to specific approach and the long run relation is as follows:

\[ m_{i,t} = 0.61 * m_{i,t-1} + 2.30 * \Delta \ln(w_i/w)_{t-1} - 0.17 * (u_i/u)_{t-1} + 0.13 \]

Hence differences in relative regional wage growth influence net migration rates!

They then estimate an adjustment equation for each region’s relative wage in terms of the unemployment differential of the region. The adjustment mechanism that they estimate is consistent with regional Phillips curve dynamics, despite the apparent disparity between Phillips curve analysis and compensating equilibrium. In the long-run they find that if a region has on average a one percentage point of unemployment above the national average, its wage rate would be 3.2% above the average wage rate for the country as a whole. Relative wages compensate for unemployment differences at the ratio 3:1. This is a compensating differential. They thus find some support for the equilibrium approach.

Their conclusion is that inter-regional migration was found to respond to changes in regional relative wages and to differences in employment opportunities as measured by regional unemployment ratios. There exists an adjustment towards a long-run equilibrium (where only compensating differentials exist) but the adjustment is slow. It should however be noted that in this study, cost of living differences among regions as represented by regional consumer or house prices (rents) are neglected, except for the part accounted for by region specific dummies.

Daveri and Faini (1999) use aggregate data from southern Italian regions to test whether risk is a significant determinant of the decision to emigrate. They focus on the choice between internal and international migration as alternative means for risk diversification. The general idea behind their paper is that migration may be seen as an opportunity to diversify risk for the family; if returns to different locations are imperfectly correlated, households could reduce total income risk by having some of their working members sent to a variety of locations. Migration may then take place even in the absence of significant wage and unemployment differentials.

The theoretical model builds on the portfolio approach to the determination of the optimal family size developed by Appelbaum and Katz (1991). They also cite Lucas (1997), advancing that both wage and expected employment opportunities are crucial factors in shaping the behaviour of potential migrants.

Data is from southern Italian regions for the 1970’s and 1980’s. The determinants of migration that are included in the regressions are a set of risk indicators and a group of control variables suggested by previous studies: per-capita
labour income, unemployment rates, share of agricultural and construction employment, index of human capital endowment, share of young people (15-29) in the population.

The estimated equation is as follows:

\[ m_j = b_0 + b_1 \ln(w) + b_2 \ln(1-u) + \sum_{j=d,f} b_{3j} \rho_j + b_4 \alpha + b_5 \times SHAG \\
+ b_6 \times SHC + b_7 \times HK + b_8 \times AGE + v_j \]

where

- \( m_j \) = number of migrants \( M_j \) from the home region to destination \( j \) (\( j= \) domestic, foreign) divided by the total population in the sending region;
- \( w \) = labor cost in the region of origin;
- \( u \) = unemployment rate in the region of origin;
- \( \rho_j \) = correlation coefficient between income at home and in region \( j \);
- \( \alpha \) = variability of income in the region of origin (coefficient of variation of GDP);
- \( SHAG \) and \( SHC \) = respectively the shares of agricultural and construction employment;
- \( HK \) = human capital (secondary enrolment rate lagged five years);
- \( AGE \) = proportion of people aged 15-29 (region of origin).

Daveri and Faini find that risk is a significant determinant of migration decisions. A rise in the correlation of southern Italy’s and foreign incomes reduces foreign emigration and increases domestic emigration; an increase in the correlation of southern and northern Italy’s incomes depresses domestic migration. Concerning domestic migration, they find that wages are a significant determinant whereas unemployment is not (except in one specification out of 5): a higher home wage is associated with lower emigration. Concerning international migrations, wages are still significant but not the employment rate; only when the coefficients on \( w \) and \( 1-u \) are constrained to be the same, does the employment rate become a significant determinant of migration. It should be noted that they do not consider questions of degree of integration of the variables entering the equations.

The work of Gorbey, James and Poot (1999) is concerned with migration between Australia and New-Zealand. It describes a hybrid methodology for forecasting the New-Zealand population two years ahead (NB hybrid means the combination of a cohort method for natural increase and a statistical method for migrations). Indeed, in the short run, migration tend to evolve faster than natural increase. This has always been an important problem in forecasting regional population change but it is becoming a problem in forecasting population change at the national level as well. Since cross-border mobility is increasing globally, migration modeling becomes a major task in population forecasting.

In the first place, a structural econometric model was successfully estimated with annual data for the period from 1953 to 1990. It comprises four equations of which explanatory variables are: real airfares, age composition of the
population, return migration, macroeconomic conditions (inflation and unemployment), growth in real earnings associated with a move, relative employment growth in the two countries). However, there are some disadvantages to the use of such a model:
- migration forecasts can only be computed after forecasts of the exogenous variables of the model have been obtained;
- the potential unstability of the behavioural relationships
- the absence of a proper theoretical background implying too much reliance on goodness-of-fit measures.

The authors then switched to VAR modeling, but here again, some problems arouse:
- too many parameters to be estimated;
- general multicollinearity.

A BVAR-approach (Bayesian VAR with restrictions on the coefficients) was chosen in order to overcome overparametrization: in fact, both BVAR and VAR models have been estimated.

The theoretical framework that guided the choice of variables is the conventional one in which migration is seen as the outcome of an investment evaluation by labour force participants individually or as households (Sjaastad 1962, Greenwood 1985):

\[ NMR_{a,t} = f(Er_{a,t}, Eu_{a,t}) \]

where
- \( NMR_{a,t} \) refers to the normalized net migration flow,
- \( Er_{a,t} \) is the expected average monetary rate of return from migration from region \( a \);
- \( Eu_{a,t} \) the expected positive or negative non-pecuniary gain from migration (=location specific amenities).

\( Er_{a,t} \) will be a function of the relative expected income adjusted for cost-of-living differentials (Greenwood 1991).

The models were estimated with quarterly data from the first quarter of 1975 until the 1st quarter of 1995. The choice of the variables was guided by the availability of quarterly data for AUS and NZL on a comparable basis and by earlier research. Unit roots tests (ADF, Phillips-Perron) had been carried out. The quarterly net trans-tasman migration rate was found to be stationary. Explanatory variables were mostly non-stationary hence the use of first differences in the regressions.

Bivariate Granger-Sims causality tests have been carried out to gauge the strength and direction of the relationships. These tests confirmed that trans-tasman migration is economically driven along the lines of the standard neoclassical model of migration. There is no evidence for a reverse causality, i.e. that migration causes GDP or the earnings ratio. Granger tests provided the basis for the candidate variables but models were selected on the basis of their out-of-sample forecasting performance (Theil U-statistic). VAR and BVAR models were estimated with seasonally unadjusted data. The forecasts of the best
ARIMA model were considerably worse than the BVAR forecasts.

In Karemera, Oguledo and Davis (2000), a modified gravity model (MGM) of migration is specified and estimated with immigration data for selected countries, examining the factors affecting migrant flows to North America (Canada and USA). The MGM is applied to data arranged in a time-series and cross-section form, 68 originating countries being distinguished, both from industrialized and from emerging economies, for the period 1976 to 1986. Separate equations are modeled for Canada and the USA. The main theoretical references cited are Borjas (1989), Greenwood (1975) and Greenwood (1991).

The following variables are included as explaining factors in the migration equation: distance, population, income, inflation, unemployment, language, business credit ratings, political instability, political rights, relative freedom, civil liberties, immigration policy, continent.

The estimated model is as follows:

$$M_{ijt} = Z_{ijt} \ast B + U_{ij} + \lambda_t + v_{ijt}$$

where

- $M_{ijt}$ = migration from i to j, at time period t;
- $Z_{ijt}$ = matrix of migrant flow determinant vectors;
- $U_{ij}$ = migrant flow effects associated with the pair of countries i and j;
- $\lambda_t$ = time effects specific to a particular year.

Results for economic variables (for those of interest for this study):
- the level of economic development (as measured by GDP per capita) is correctly signed, i.e. positive for USA and CAN but not always significant; in the case of the USA equation it is also (correctly) negative for sending country but this is not true for the Canadian model;
- inflation has the correct negative sign for receiving countries (though not significant) but not for the sending countries;
- unemployment is not a determinant factor: it is either wrong signed (Canada) or it is not significant (USA, sending countries).

The objective of Parikh and Van Leuvensteijn (2003) is to study the determinants of labour migration in Germany using data on unemployment and wage differences between source and destination regions. The theoretical foundation is the "option value approach" to migration as put forward in Dixit and Pindyck (1994), Burda (1993) and Burda et al. (1998). According to those, the opportunity cost of migrating today, in addition to the expected net present value of future income gains from migration net of migration costs, is referred to as the option value of waiting. The option value of waiting is calculated as the difference between expected net-present value from postponing migration and the expected net-present value from migrating today. Even if the expected net-present value is positive, the prospective migrant may not migrate if the fixed cost of migrating is sufficiently high. Such fixed costs could include pecuniary components associated with physically moving a household from one place to another. If risk aversion were to be introduced in this model, it could change the migration decision. In the model, the option value of waiting is negatively
related to the discount rate and positively to the rate of wage convergence between regions. If wage convergence exists, waiting to move may increase the worker’s wage without any costs of migration. Waiting is also sensible when the migration costs diminish or the value of time of a person decreases. This paper argues that under wage convergence, between regions of Germany, labour mobility could be low despite high levels of prevailing unemployment.

The study uses data on interregional labour migration (as opposite to population migration) for the years 1993-1995, for the 16 Länder of Germany (both eastern and western). Data on regional prices were not available. The study resorted to remove endogeneity as wage differences and unemployment differences could be determined by migration rather than migration being determined by these two. The former was removed by using instrumental variable estimation where previous years’ similar values were used such as wage and unemployment differences of 1992 for 1993 and similarly 1993 for 1994 and 1994 for 1995.

The model estimated is as follows:

$$\ln \left( \frac{L_{ijt}}{L_{it} * L_{jt}} \right) = \beta_1 + \beta_2 * LUNEM_{it} + \beta_3 * UR_{ijt} + \beta_4 * WBLUET_{ijt} + \beta_5 * WWITET_{ijt} + \beta_6 * OWNED_{ijt} + \nu_{ijt}$$

where $t = 1993, 1994, 1995$ and

$L_{ijt} =$ gross labour flows from $i$ to $j$;
$L_{it} =$ labour stocks;
$L_{ijt}/(L_{it} * L_{jt}) =$ normalized labour flow;
$LUNEM_{it} =$ level of unemployment in each region;
$UR_{ijt} =$ unemployment rate differentials;
$WBLUET_{ijt}$ and $WWITET_{ijt} =$ differences in wages;
$OWNED_{ijt} =$ differences in house ownership.

Main conclusions are:

- the fixed effects model dominates the random effects model in each case and confirms the hypothesis of a U-shaped relationship between white-collar workers’ wage differentials of receiving and sending region and migration;
- unemployment level and differences in unemployment between receiving and sending regions was never significant; it appeared that when wage convergence was rapid prospective immigrants tended to wait since the opportunity cost of migration was rising;
- the relationship between income inequality and migration did not turn out to be strong;
- there exists weak evidence that immigration is positively related with inequality among white-collar worker’s wages while it is negatively related to income inequality of blue-collar workers’ wages;
- one important limitation of the study is the absence of information on regional cost of living.

Ashby (2007) tries to determine whether individuals in aggregate migrate toward states with higher economic freedom. His study is indeed centered on the
analysis of the impact of government policies that enhance economic freedom on gross migration flows in the lower 48 US-States.

Under certain assumptions, consumers convey their preferences through migration or "voting with their feet". If a country meets the standards as a "market-preserving federation", migration is likely to result in more efficient policies that result in greater well-being of individuals in the long-run. In other words: the more economic authority is decentralized, the more efficient the economic outcome will be. But when local governments systematically receive economic assistance from higher levels of government, this violates the principle of market-preserving federalism, endangering the sustainability of markets and opportunities for efficient utility maximization through migration. Still in other words: if individuals are free to migrate between jurisdictions, this serves as a check on government intrusion in economic matters.

The index of economic freedom that is used in this study is made up of three major components:

- size of government;
- takings (i.e. general tax burden) and discriminatory taxation;
- labour market freedom.

Economic freedom thus takes into account government consumption, transfers, social security, tax revenues, marginal tax rates, minimum wages, government employment and union density.

Ashby (2007) uses a form of the modified gravity model which diverges in several aspects:

- the use of contiguity matrices, to take into account spatial proximity;
- the use of a spatial autoregressive estimator in order to test for spatial effects;
- the fact that the dependent variable is the "log-odds" of the migration flows (cf. below) as opposed to the classical gross or net migration rate.

The model that is estimated is the following:

\[
\ln \left( \frac{m_{ij}}{1 - m_{ij}} \right) = \alpha_0 * Stay_{ij} + \beta_0 * Move_{ij} + \rho W * \ln \left( \frac{m_{ij}}{1 - m_{ij}} \right) + \\
\beta_1 * \frac{EcFree_j}{EcFree_i} + \beta_2 * \ln \left( \frac{Pop_j}{Pop_i} \right) + \beta_3 * \ln \left( \frac{Density_j}{Density_i} \right) + \\
\beta_4 * \ln \left( \frac{Income_j}{Income_i} \right) + \beta_5 * \left( \frac{EmplGrowth_j}{EmplGrowth_i} \right) + \\
\beta_6 * \left( \frac{Retired_j}{Retired_i} \right) + \beta_7 * \left( \frac{HeatDays_j}{HeatDays_i} \right) + \\
\beta_8 * \left( \frac{Precipitation_j}{Precipitation_i} \right) + \delta_1 * Dis_{ij} + \delta_2 * Dis_{ij}^2 + \\
\lambda W_{ij} + \varepsilon_{ij}
\]

Explanations:
ln \left( \frac{m_{ij}}{1-m_{ij}} \right) \text{ is the log-odds ratio of migrations: } m_{ij} \text{ are gross migration rates, i.e. the amount of migrants from the origin to the destination divided by the origin population at the beginning of the period; the formula centers the dependent variable around zero;}

Stay_{ij} = \text{ dummy variable equal to 1 for all stayers (i.e. where i=j)};
Move_{ij} = \text{ dummy variable equal to 1 for all movers (i.e. where migration occurred)};

W = \text{ contiguity matrix that measures proximity to a given observation/state: each individual row of this matrix represents the contiguous relationship of a destination state; if this particular destination state has a contiguous relationship with the origin state, this will positively impact migration to this state (} \rho > 0);\n\rho \text{ measures the effect that neighbouring observations have on the dependent variable; } \lambda \text{ measures correlation in the error term: if } \rho \text{ and/or } \lambda \text{ are different from zero, then spatial dependence exists;}

Density_j/Density_i = \text{ relative population density;}
Income_j/Income_i = \text{ real mean income per capita averages between 1990 and 1995;}
EmplGrowth_j/EmplGrowth_i = \text{ difference in employment growth rates;}
Retired_j/Retired_i = \text{ persons over 65;}
Dis_{ij} = \text{ proxy for travel costs.}

The explanatory variables are lagged since the driving forces of migration are supposed to act with delay.

Main results (in the interest of this study) are that relative income and relative employment growth are significant independent variables (as are most of the other candidate variables). Economic freedom is an explaining variable only if income and employment growth are taken from the regression.

Mathä & Wintr (2007) investigate the aggregate cross-border commuting pattern in the bordering regions of four EU countries: Belgium, France, Germany and Luxembourg. They estimate an extended (or modified) gravity model to explain commuting flows between the four countries, for the period 1996-2003, in the bordering regions of Luxembourg. The first estimated model is as follows:

$$\ln(T_{ijt}) = \alpha \ln(L_{it}) + \beta \ln(E_{jt}) + \gamma_1 d_{ij} + \gamma_2 d_{ij}^2 + \delta d_{lang_{ij}} + \sum_r \nu_r d_{reg_{ijr}} + \varepsilon_{ijt}$$

where

T_{ij} = \text{ commuting flows;}
L_{it} = \text{ size of labour force in home region;}
E_{jt} = \text{ employment in the region of work;}
d_{ij} = \text{ distance between the regions;}
d_{lang_{ij}} = \text{ language dummies for common language;}
d_{reg_{ijr}} = \text{ dummy variables for each region.}
All variables in this model are significant and the it explains 97% of the variability of the dependent variable. In a second model where unemployment differences and per capita gross value added (to proxy revenues) are added, population and labour force size variables turn insignificant. Unemployment differentials are insignificant too. According to the authors, this result is due to a high collinearity between the relative unemployment rate and gross value added per employee.

Finally, the authors estimate a negative binomial model and a generalized negative binomial model where the probability of $T_{ijt}$ people commuting from $i$ to $j$ is drawn from a Poisson distribution with parameter $\lambda_{ijt}$ (the mean). The results of the negative binomial model reveal a smaller elasticity of commuter flows with respect to relative gross value added per employee and employment in the region of work. In the generalized negative binomial model, which performs slightly better, all variables have the expected sign and are significant at the 1% level except employment in the region of destination.

Marvakov and Mathä (2007) argue that regional labour mobility is of increasing concern in the context of the Single European Monetary Policy, as EMU implies a reduction of national policy options. Thus, it would be important that the remaining adjustment mechanisms function effectively. While most of the empirical literature focuses on labour mobility in terms of migration, their paper provides an empirical assessment of the determinants of aggregate regional commuting flows in the EU, an issue often examined in a local or national context but still un(der)explored on EU level. Using an extended gravity framework, commuting is found to respond to differences in regional wages and unemployment, and to provide an equilibrating mechanism to labour market disequilibria. Higher levels of education and labour force participation of women, as well as a larger services sector are associated with a higher percentage of commuting. Finally, the results reveal interesting geographical differences between internal, border and coastal regions.

9.3 Tests for the order of integration of the series entering the regressions

Unit root tests are a form of pre-testing, ideally (although not necessarily) to be carried out before cointegration analysis and regressions. A proper long-run specification as the one laid out in (5) optimally consists of series that are all I(1).

Maddala and Kim (2002) strongly advise not to use the common augmented Dickey-Fuller test (ADF) but admit that it is still widely applied in empirical work. They argue in favour of the DF-GLS test, as brought forward by Elliott, Rothenberg and Stock (1996). Both have been carried out for this study, results are in table xy. Recall that the sample is 1980-2007 with yearly data.

For the ADF test, the option "constant and no time trend" has been used, together with automatic lag length selection (Akaike criterion, integrated in Eviews). Significance levels are according to McKinnon (1996) as included in the Eviews Software package.
Results indicate that all candidate series to be used in the estimation\textsuperscript{11} of (5) clearly seem to be I(1) except one, namely the ratio of cross-border workers to domestic employment, even though the DF-GLS test gives some weak evidence in that direction.

\textsuperscript{11}Series entering the regressions are represented on a yellow (grey) background.
9.4 Data and results
Figures 9-14: Data used in the regressions (2/2)

Fig. 9: Net wages, by country

Fig. 10: Net wages

Fig. 11: Regional prices for construction land

Fig. 12: Regional house prices

Fig. 13: House prices

Fig. 14: Independent variables, ratios (Luxbg / other)
Table 1: unit root tests

<table>
<thead>
<tr>
<th>Tested series (1980-2007, yearly)</th>
<th>ADF</th>
<th>DF GLS</th>
<th>ADF</th>
<th>DF GLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immigration</td>
<td>-0.20</td>
<td>-0.70</td>
<td>-3.99</td>
<td>-3.46</td>
</tr>
<tr>
<td>Migrant</td>
<td>-0.78</td>
<td>-0.73</td>
<td>-5.94</td>
<td>-5.44</td>
</tr>
<tr>
<td>Total population</td>
<td>0.39</td>
<td>-1.63</td>
<td>-2.53</td>
<td>-0.98</td>
</tr>
<tr>
<td>Immigration rate</td>
<td>-0.71</td>
<td>-0.55</td>
<td>-3.96</td>
<td>-3.49</td>
</tr>
<tr>
<td>Migrant / tot. pop.</td>
<td>-2.92</td>
<td>-2.97</td>
<td>-6.06</td>
<td>-5.52</td>
</tr>
<tr>
<td>Cross-border workers (CBW)</td>
<td>-1.57</td>
<td>-0.84</td>
<td>-2.79</td>
<td>-2.33</td>
</tr>
<tr>
<td>Total employment</td>
<td>1.20</td>
<td>-0.96</td>
<td>-3.20</td>
<td>-2.51</td>
</tr>
<tr>
<td>Cross-border ratio 1</td>
<td>-1.13</td>
<td>-0.90</td>
<td>-2.56</td>
<td>-2.30</td>
</tr>
<tr>
<td>Cross-border ratio 2</td>
<td>-2.82 *</td>
<td>-1.35</td>
<td>-2.16</td>
<td>-1.98 *</td>
</tr>
<tr>
<td>House prices Lux</td>
<td>0.52</td>
<td>0.22</td>
<td>-2.15</td>
<td>-2.16 **</td>
</tr>
<tr>
<td>House prices abroad</td>
<td>-0.53</td>
<td>0.17</td>
<td>-4.07</td>
<td>-4.16 ***</td>
</tr>
<tr>
<td>Ratio of house prices</td>
<td>0.98</td>
<td>0.29</td>
<td>-3.59</td>
<td>-3.64</td>
</tr>
<tr>
<td>Net real income Lux</td>
<td>0.15</td>
<td>-0.23</td>
<td>-4.67</td>
<td>-3.72</td>
</tr>
<tr>
<td>Net real income abroad</td>
<td>-0.19</td>
<td>-1.19</td>
<td>-4.42</td>
<td>-4.26 **</td>
</tr>
<tr>
<td>Income ratio</td>
<td>-1.32</td>
<td>-0.65</td>
<td>-5.43</td>
<td>-5.33 ***</td>
</tr>
<tr>
<td>Resident unempl. rate</td>
<td>-2.53</td>
<td>-2.65 **</td>
<td>-3.03</td>
<td>-3.12 **</td>
</tr>
<tr>
<td>Foreign unempl. rate</td>
<td>-5.99 ***</td>
<td>-3.01 ***</td>
<td>-3.15 **</td>
<td>-2.45 **</td>
</tr>
<tr>
<td>Ratio of UE rates</td>
<td>0.47</td>
<td>-1.09</td>
<td>-2.46</td>
<td>-2.41 **</td>
</tr>
</tbody>
</table>

Significance levels according to McKinnon (1996): "Fractional distribution functions for unit root and cointegration tests" (Journal of Applied Econometrics, 11, 601-8), included in the Enasov Software package.

Table 2: OLS estimation of long-run relations; ADF test on residuals

<table>
<thead>
<tr>
<th>Dependent</th>
<th>Independent</th>
<th>ADF-test*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employ / population</td>
<td>Unemployment rates</td>
<td>Net wages</td>
</tr>
<tr>
<td>LIP / RUP</td>
<td>UR / RUP</td>
<td>MIP / NIP</td>
</tr>
<tr>
<td>Immigration ratio</td>
<td>MIP / P</td>
<td>0.59</td>
</tr>
<tr>
<td>MIP / P</td>
<td>1.69</td>
<td>-0.19</td>
</tr>
<tr>
<td>Outmigration ratio</td>
<td>MIP / P</td>
<td>-0.05</td>
</tr>
<tr>
<td>MIP / P</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cross-border workers</td>
<td>MIP / P</td>
<td>3.59</td>
</tr>
<tr>
<td>MIP / P</td>
<td>1.68</td>
<td>-0.17</td>
</tr>
</tbody>
</table>

* Residual based co-integration test; cf. McKinnon (1991) as cited by MK (2002); critical value = approx. -4.16 at the 5% level (cf. MK 2002, p. 201)

** Constrained
Table 3: ECM results (elasticities*)

| Dependent var | Legal immigration | Legal emigration | Emigration | Real Household | Real wage abroad | Real Household | Real wage abroad | Real household prices | Real wage abroad | Durbin | Value | Std dev | Adjusted R-squared | S.E. Regression | F-statistic | Joint Ml |
|---------------|------------------|------------------|------------|----------------|-----------------|----------------|-----------------|-----------------------|-----------------|--------|-------|--------|---------|-----------------|----------------|------------|--------|
| Immigration   | 6.29             | n.a.             | 1.39       | 0.41           | -0.07           | n.a.           | 0.59            | 0.79                  | -0.08           | 0.48   | 2.75  | 0.25   | 0.059  | 0.10            | 0.06          |
| Emigration    | 0.24             | n.a.             | -0.83     | 0.14           | 0.31            | n.a.           | -0.89           | 0.46                  | -0.25           | -1.51  | 0.48  | 0.059 | 0.50    | 0.84          |
| Cross-border workers  | n.a.             | n.a.             | 1.22     | 0.21           | -0.08           | n.a.           | 0.04            | 1.00                   | 1.54           | -1.20  | 1.51  | 0.34   | 0.010  | 0.58          |

* All variables in log-form

1 All migration variables are expressed as migration rate, i.e. migration flow divided by total population, the stock of cross-border workers is divided by total employment.
2 Coefficients in bold are elasticities; **= 10% significance level; ***= 5% significance level; **= 1% significance level; n.s. = not significant (n.s.) at the 10% level (short run) except for the first two equations where the long-run part is calibrated.
3 Real wage abroad
4 Real household prices abroad
5 House prices abroad
6 Real wage abroad
7 Coefficients in italics are constrained to 1.
8 Coefficients in italics are constrained to 1.
9 n.a. = not applicable, i.e. not included in the regression.
Figures 20-23: decrease of banking sector activity, full model results vs. exogenous migrations (2/2)
Figures 24-30: shocks on three exogenous variables governing foreign labour supply

Fig. 24: GDP vol.

Fig. 25: domestic demand

Fig. 26: real net disposable income

Fig. 27: resident unemployment rate

Fig. 28: total employment

Fig. 29: cross-border workers

Fig. 30: resident employment
Figures 31-32: further analysis of the shock on foreign revenues

Fig. 31: shock on foreign net wages - 1

Fig. 32: Shock on foreign net wages - 2