Estimating and forecasting European migration: methods, problems and results[1]

Herbert Brücker and Boriss Siliverstovs

The specification of macro migration models and, hence, forecasts of migration potentials differ largely in the literature. Two main differences characterise macro migration models: first, whether migration flows or stocks are used as the dependent variable, and, second, whether the heterogeneity in the migration behaviour across countries is considered. This paper addresses both issues empirically using German migration data from 18 European source countries in the period 1967–2001. It finds first that panel unit-root and cointegration tests reject the hypothesis that the variables of the flow model form a cointegrated set, while the hypothesis of cointegration is not rejected for the stock model. The second finding is that standard fixed effects estimators dominate the forecasting performance of both pooled OLS and heterogeneous estimators. Applying the preferred fixed effects estimator, the migration potential from the Central and Eastern European accession countries is estimated at 2.3–2.5 million persons for Germany, which implies a migration potential of 3.8–3.9 million persons for the EU-15. Finally, our estimates indicate that the migration potential in the EU-15 is already exhausted and that the migration potential from Turkey is relatively small.

Contents

[1] This paper benefited substantially from comments and suggestions by Michael Fertig, the editor Elmar Hönekopp and two anonymous referees. The authors alone are however responsible for any remaining errors.

1 Introduction

International migration is the "great absentee" (Faini *et al.* 1999) in the liberalisation of global goods and factor markets. While the barriers to international trade and capital mobility are already largely removed, the opening of labour markets lags far behind. Moreover, most regional trade areas in the world exclude labour markets from the removal of barriers to trade and factor movements. The European Union (EU) forms a notable exception in this context. The free mobility of labour and other persons is defined as one of the four fundamental freedoms of the Common Market since the Treaty of Rome, which established the Community in 1957. The free movement started in a community of six countries with a joint population of 185 millions in 1968, and has been step by step extended to the EU-15 and the three other members of the European Economic Association (EEA) with a joint population of 380 million persons until the mid of the 1990s. Another eight countries from Central and Eastern Europe together with Cyprus and Malta have become members of the EU in May 2004. Moreover, Bulgaria and Romania will probably join the Community in 2007. Altogether, the current extension of the EU to the East will increase its population by some 102 million persons. However, transitional periods for the free movement of labour have been agreed which allow the individual Member States to dispense free labour mobility up to a maximum period of seven years.

At least from a global perspective, the EU and the EEA formed a club of rich countries with relative homogeneous per capita income levels in the past. Even in the case of the EU's Southern enlargement, the per capita income of Greece, Portugal and Spain have already converged to 65–70 % of the average level of the old Member States, when EU membership was granted (Maddison 1995). Consequently, less than one-third of the foreign population in the EU stems from other Member States. The main source of migrants are middle- and low income countries in the neighbouring regions of the EU, i.e. Northern Africa (Algeria, Morocco, Tunisia) and South-eastern Europe (Turkey and the Balkan countries).

The current enlargement round changes the picture of the EU as a club of rich countries with homogeneous income levels. The average GDP per capita measured at purchasing power parities (PPP) of the eight accession countries from Central and Eastern Europe which are admitted in this enlargement round is estimated by Eurostat (2003) at almost 50 per cent of that in the EU-15. If Bulgaria and Romania join the EU, the average PPP-GDP per capita of the accession countries will decline to 45 per cent of the EU-15 level. At current exchange rates, the per capita GDP of the accession countries is, at 20 per cent of that in the EU-15, even lower. This income gap is larger than in any other of the previous enlargement rounds of the Community. Nevertheless, the accession countries from Central and Eastern Europe are middle income countries from a global perspective, and their income levels are twice as high as that of the remaining neighbours of the EU in Northern Africa, South-eastern Europe and the Commonwealth of Independent States (CIS) (see Figure 1).

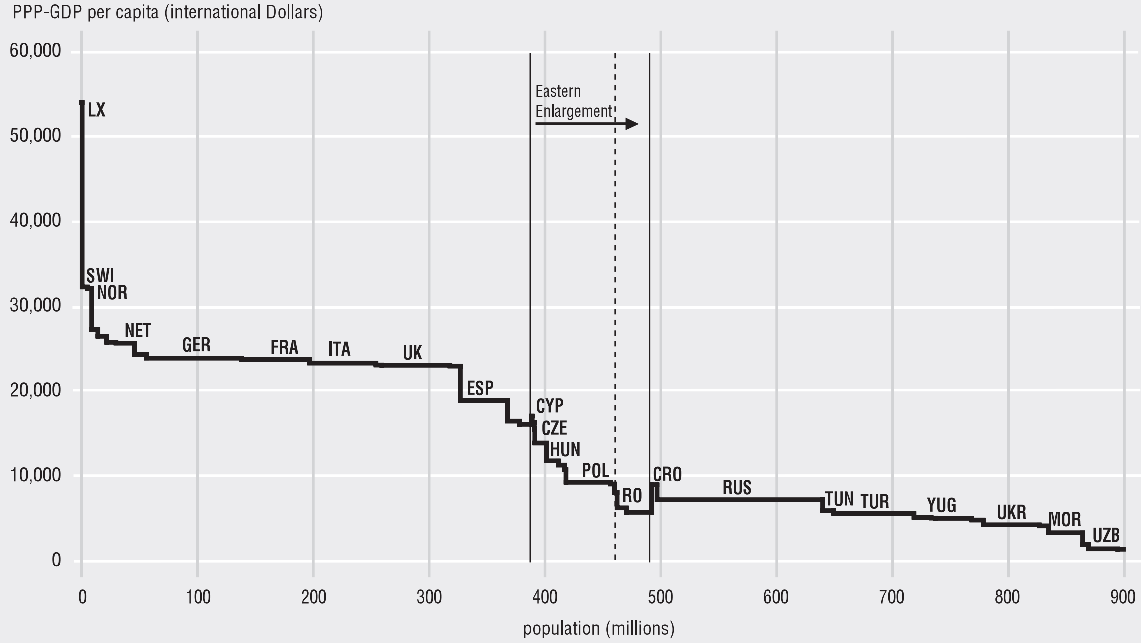
Given the unprecedented income differences between the old and the new Member States of the EU, the uncertainty on the migration potential is large. Starting with the seminal contribution of Layard *et al.* (1992), numerous studies have tried to reduce this uncertainty. Basically we can distinguish three approaches to estimate the migration potential from the East in the literature: representative surveys, extrapolations of South-North migration to East-West migration, and forecasts based on econometric studies.

Representative surveys of the population are affected by three problems, which make it hard to draw quantitative inferences from them. First, it is difficult to assess the extent to which the migration intentions revealed in surveys later materialise into actual movements of individuals or households. Second, surveys capture only the supply side and ignore demand-side factors such as job opportunities and the availability of housing. Third, surveys cannot appropriately capture the temporary dimension of migration. Since only a minority of migrants stay in a foreign country permanently, a large number of individuals who will migrate at a certain point of their life may coexist with a small number of persons who are living abroad at a particular point in time. As a consequence, findings from surveys of the population from the accession countries vary widely: individuals intending to migrate range from 2–30 % of the population, depending on the questionnaires employed and the interpretation of the answers.[2] Note that the correlation between migration intentions and actual migration is rather weak: for example, according to the German Socio-Economic Panel (GSOEP), more than 10 % of the East German pop-

[2] The careful studies by Fassmann/Hintermann (1997), IOM (1998), and Krieger (2003) obtain results at the lower end of this spectrum. However, their results depend critically on the interpretation and weighting of the answers.

Figure 1

The European income gap, 2002



Source: World Bank (2004), author's calculations.

ulation intended to migrate to Western Germany in 1991, but only 5 % of those who expressed the intention to migrate had actually moved five years later.

Several studies have extrapolated the number of South-North migrants in the 1960s and early 1970s to East-West migration. The income gap between the Southern and the Northern European countries in the 1960s was similar to the gap between the EU-15 and the accession countries today. Moreover, although the Southern European countries were not EU Members at that time, bilateral guestworker agreements led to the *de facto* opening of Northern European labour markets and supported labour migration until the first oil price shock of 1973. In general, these extrapolation studies find a long-run migration potential of around 3 % of the population (e.g. Layard *et al*. 1992). However, in stark contrast to the conditions for South-North migration in the early 1960s and 1970s, the conditions for East-West migration today are affected by imbalances in both the labour markets of the receiving and sending countries, incomplete recovery from the transition shock, and close geographical proximity. Thus, extrapolation studies can provide no more than a hint at plausible orders of magnitude.

The majority of the forecasts of East-West migration are based on econometric estimates, which usually explain migration flows or stocks by variables such as the income differential, (un-)employment rates, and some institutional variables. Although most studies employ the same set of explanatory variables, the estimates of the parameters, and, hence, of migration potentials differ considerably in the literature. Table 1 presents an overview[3] on a number of econometric studies, which have tried to estimate the migration potential from the East either for Germany or the EU-15.[4] Since the forecasts refer to different Central and Eastern European countries, their findings have been expressed here as a percentage of the population in the sending countries.

As can be seen in Table 1, the forecasts for the initial net migration range from 0.02 – 0.64 % of the population in the sending countries for Germany, which corresponds to a figure of 20,000 – 640,000 persons per annum. Analogously, the results for the long-run migration stock range from 2.3 – 7.2 % for Germany, or, if we assume that the present regional distribution of migrants from the accession countries remains constant, from 3.8 – 12.0 % for the EU-15. Given the policy relevance of the issue, it is worth-

[3] This overview is far from complete, for surveys of the literature see Hönekopp (2000) and Straubhaar (2002).
[4] Note that the focus on Germany in these studies is not accidental, since Germany absorbs around 60 % of the migrants from the Central and Eastern European accession countries in the EU-15 (Alvarez-Plata *et al*. 2003).

Table 1

Econometric estimates of the migration potential from the Central and Eastern European accession countries

study	data base	model	Estimation method	projection results	
				Initial net flows	long-run stocks
Alvarez-Plata et al. (2003)	migration from 18 countries to Germany, 1967–2001	dynamic stock model	fixed effects	0.22% (GER) 0.33% (EU15)	2.33% (GER) 3.82% (EU15)
Boeri/Brücker (2001), Brücker (2001)	migration from 18 countries to Germany, 1967–98	dynamic stock model	fixed effects	0.22% (GER) 0.34% (EU15)	2.53% (GER) 3.89% (EU15)
Dustmann et al. (2003)	migration from 17 countries to Germany and UK, 1960–94	static flow model	GMM	0.02–0.2% (GER) 0.004–0.01% (UK)	– –
Flaig (2001), Sinn et al. (2001)	migration from 5 countries to Germany, 1974–97	dynamic stock model	pooled OLS	0.64% (GER)	7.20% (GER)
Fertig (2001)	migration from 17 countries to Germany, 1960–97	dynamic flow model	fixed effects	0.07% (GER)	–
Fertig/Schmidt (2001)	migration from 17 countries to Germany, 1960-97	static flow model	GMM	0.01–0.06% (GER)	–
Hille/Straubhaar (2001), Straubhaar (2001)	migration from 3 source to 8 receiving countries, 1988–99	static flow model	pooled OLS	0.27% (EU)	

while to study the causes of these differences in depth.

Beyond different data sources, two main aspects distinguishes the approaches in the econometric forecasts: The first difference refers to the choice of the dependent variable. The major part of the studies follows the standard approach in the literature and use migration flows as the dependent variable. Since forecasts of the additional labour supply through migration can hardly be based on gross inflows, net migration rates are usually employed. In contrast, another part of the studies use migration stocks as the dependent variable. The choice of the dependent variable has important consequences for modelling the dynamics of the migration process. The flow model implicitly assumes that migration will continue until (expected) income levels converge to a certain threshold level, where the costs of migration exceeds the benefits, while the stock model predicts that net migration will eventually come to a halt even if large income differentials persist. The two models have different micro foundations. While the flow model relies on the concept of a representative agent, the stock model is based on the assumption that preferences or human capital characteristics of individuals differ, such that the benefits and costs of migration are not equal across individuals.

The question whether stock or flow models are adequate to model macro migration functions is not purely of academic interest. In the case of Southern enlargement, the introduction of free movement has not resulted in increasing migration stocks, although the income of the Southern Member States amounted to no more than 65 – 70 % of the EU average at that time. In terms of the stock model, this phenomenon can be explained by the fact that migration stocks had already achieved their equilibrium levels when the free movement was introduced.

The second important difference between the studies refers to the estimation method. It is uncontroversial that country-specific factors such as culture, language, history, geography, etc. affect the benefits and costs of migration, and consequently, the migration propensity across countries. These factors are only partly observable and can therefore not be included completely in migration models. The heterogeneity across countries is however only partially considered by the migration models presented in Table 1 – if at all. A number of studies apply pooled OLS estimators, assuming that both the intercept and the slope parameters are homogenous across countries. Another part of the studies use fixed effects models, assuming that the intercept differs across countries, while the slope parameters are homogenous. For the out-of-country forecasts the fixed effects are explained in an auxiliary regression by time-invariant variables (e.g. Boeri/Brücker 2001; Fertig 2001). The quantitative results of these two approaches differ considerably. As an example, while Boeri/Brücker (2001), Brücker (2001) and the follow-up study by Alvarez-Plata *et al*. (2003) forecast for Germany a long-run migration potential of 2.3–2.5 % of the population from the accession countries using a fixed effects estimator, Sinn *et al*. (2001) and Flaig (2001) estimate the long-run migration potential at 7.2 % on basis of a pooled OLS model.

It is well-known that pooled OLS models can yield biased and inconsistent results if omitted variables are correlated with the explanatory variables (Baltagi 1995). Fixed effects models avoid this estimation bias, but can still result in biased and inconsistent estimates if the slope parameters differ across countries (Pesaran/Smith 1995). Moreover, with dynamic fixed effects models, simultaneous equation bias can affect the estimation results. As an alternative, heterogeneous estimators can be applied. These heterogeneous estimators are based on individual regressions. For out-of-country forecasts, averages of the estimated parameters can be used. However, in samples with a limited group and time dimension, heterogeneous estimators can yield unstable results, such that homogenous panel estimators might outperform heterogeneous estimators (see e.g. Baltagi *et al*. 2000).

Thus, both the adequate specification of macro migration models and the choice of the appropriate estimation procedure are controversial. Although some of these issues have been already discussed (see e.g. Alecke *et al*. 2001; Fertig/Schmidt 2001; Straubhaar 2002), a systematic analysis of these issues is missing in the literature. Against this background, this paper pursues three objectives:

- First, to examine the appropriate specification of macro migration models. More specifically, it is tested whether the standard hypothesis of macro migration models that a long-run equilibrium relationship between migration flows and the explanatory variables emerges is supported by our data. Alternatively, we consider the hypothesis that a long-run equilibrium relationship between migration stocks and the explanatory variables exists.

- Second, to analyse the quantitative consequences of different estimation methods and to compare the out-of-sample forecasting performance of different estimators in order to derive criteria for the choice of the adequate estimation procedure. To this end, we employ a wide range of estimators which are discussed in the econometric literature.

- Third, to forecast the migration potential from the accession countries on basis of the analysis carried-out before. The empirical analysis of the paper is based on migration to Germany from a panel of 18 European source countries in the period 1967–2001.

The remainder of the paper is organised as follows: Section 2 discusses alternative specifications of macro migration models and tests whether the variables of the flow or the stock model form a cointegrated set. In Section 3 we estimate the cointegrating vectors and the short run dynamics of the succeeding stock model and test the out-of-sample forecasting performance of a broad range of alternative estimation procedures. On basis of the preferred estimation procedure, Section 4 provides a projection of the migration potential from the accession countries to the EU. Finally, Section 5 concludes and discusses the policy implications of our findings.

2 Stock versus flow models

The standard macro migration model in the empirical literature is based on the fundamental assumption of a representative agent, which compares utility differences between different locations. Consequently, these models presume that a long-run equilibrium relationship between migration flows and the explanatory variables emerges. Following this line of reasoning, the long-run migration function can be written in general form as

$$m_{it} = \sum_p \beta_p X_{pft} + \sum_q \gamma_q X_{qit} + \delta\, mst_{i,t-1} + \mu_i + \varepsilon_{it}, \quad (1)$$

where $i = 1 \ldots N$ and $t = 1 \ldots T$ are the (source) country and time indices, the subscript f denotes the destination country, m_{it} is an appropriate measure for the aggregate gross or net migration rate (i.e. the number of migrants as a proportion of the population at the origin), X_{pft} and X_{qit} are vectors of observable and time-varying characteristics in the receiving country (index p) and the sending country (index q), respectively, and $mst_{i,t-1}$ is the lagged migration stock. β_p, γ_q, and δ are (vectors of) unknown parameters, μ_i captures all unobservable variables which are specific to country i and time-invariant, and ε_{it} is the error term assumed to be *iid* with zero mean and constant variance. Some models in the literature treat the country specific effects as random or assume that the intercept term is uniform across countries, which allows to include time-invariant variables such as geographical or linguistic distance. Examples for this type of specification of the macro migration function in the literature are Fields (1978), Lundborg (1991), Fertig/Schmidt (2001) and Pederson *et al*. (2004).

Most models in the empirical literature consider income variables and employment as the main explanatory variables in the specification of the empirical model, such that estimation equation has the following form (e.g. Hatton 1995):

$$m_{it} = a_1 \ln(w_{ft}/w_{it}) + a_2 \ln(w_{it}) + a_3 \ln(e_{ft}) + a_4 \ln(e_{it}) + a_5 mst_{i,t-1} + \mu_i + \varepsilon_{it}, \quad (2)$$

where w_{ft} and w_{it} are the wage rates in the receiving and the source country, respectively, and e_f and e_i are the employment rates in the receiving and source country, respectively. In addition, many empirical models include deterministic time-trends, which should capture falling transport and communication costs, and dummy variables for the institutional and political migration conditions.

This parsimonious specification of macro migration models has a long tradition in the literature. The choice of explanatory variables is primarily based on the classical contributions of Ravenstein (1889), Hicks (1932), Sjaastad (1962) and Harris and Todaro (1970). More specifically, the standard model is derived from the following assumptions: the utility of individuals is inter alia determined by expectations on income levels in the respective locations. Utility is convex in the income differential, i.e. it is implicitly assumed that other, non-pecuniary arguments enter the utility function as well. Expectations on income levels are conditioned by employment opportunities. Individuals are risk averse, but uncertainty focuses on employment opportunities. As a consequence, the model expects that the coefficient for the employment variables is larger than that for the income variables. Moreover, since employment opportunities of migrants in host countries are below those from natives, it is expected that the coefficient for the employment rate in the host country is larger than that in the source country. If capital markets are not perfect, liquidity constraints affect migration decisions. Consequently, for a given income difference between the host and the source country, the income level in the source country has a positive impact on migration (Faini/Venturini 1995; Faini/Daveri 1999). Finally, it is assumed that migration networks alleviate the costs of adapting to an unfamiliar environment, such that the costs from migration are expected to decline with the stock of migrants already existing in the host country (Massey *et al*. 1984; Massey/Espana 1987). Depending on the assumptions on the utility function, the functional form of the macro-migration function is specified both in semi-log (e.g. Hatton, 1995) and in double-log form (e.g. Lundborg 1991; Faini/Venturini 1995; Pederson *et al.* 2004).

There exist numerous micro-economic models of the migration decision in the literature which go far beyond these considerations. Inter alia, these models analyse the role of portfolio diversification of families in the absence of perfect capital markets (Stark 1991), the role of relative deprivation (Stark 1984), and the impact of uncertainty about future wage and employment conditions on the migration decision in the presence of fixed migration costs (Burda 1995; Bauer 1995). However, few of these theoretical contributions have developed macro migration functions which can be applied empirically. Moreover, the estimation of more complex macro models is hindered by data limitations. As an example, time series information on variables such as the income distribution in the receiving and sending countries is rarely available for longer time spans, which hinders the macro analysis of e.g. the role of risk aversion, risk pooling in households and the relative deprivation in migration decisions.

Thus, although the choice of variables and the functional form of the model in equation (2) relies on a number of arbitrary assumptions, there exist few alternatives to this macro migration function in the empirical literature. In this paper, however, one fundamental assumption of the standard model is called into question: as shown above, the standard model assumes that a log-linear relationship between migration flows and the economic variables exists in the long-run equilibrium. This implies that migration ceases not before (expected) income levels between the host and the source country, as determined by the wage and the employment variables on the right hand side of equation (2), have converged to a certain threshold level, which is determined by the costs of migration. In case of persistent differences in (expected) income levels, either the total population will eventually migrate or migration will not happen at all from the beginning. Note that this is a consequence of deriving macro migration functions from the concept of a representative agent, i.e. of assuming that individuals are homogeneous.

Consider as an alternative that agents are heterogeneous with respect to their preferences such that the costs of living abroad differ across individuals. Depending on their preferences, some individuals will stay at home, migrate temporarily and permanently at a given income differential. Moreover, the length of migration differs across temporary migrants. This has important consequences for the mechanics of migration stocks and flows: In the long-run, an equilibrium between migration stocks and the (expected) income differential emerges, while the net migration rate ceases to zero- at least if we assume that population growth rates are equal in the home and the migrant population. However, gross and return migration remains a positive function of the income differential in the long-run equilibrium as a consequence of temporary migration (Brücker/Schröder 2005).

Thus, under the assumption that individuals are heterogeneous, an equilibrium between migration stocks instead of migration flows and the difference in (expected) income levels in the respective locations emerges. We conceive therefore here as an alternative to the standard model in equation (2), the following specification for the long-run migration function:

$$mst_{it} = b_1 \ln(w_{ft}/w_{it}) + b_2 \ln(w_{it}) + b_3 \ln(e_{ft}) + b_4 \ln(e_{it}) + \mu_i + \varepsilon_{it}. \qquad (3)$$

The estimation of the migration functions in (2) and (3) can be affected by spurious correlation effects, if the regressions involve non-stationary variables (Granger/Newbold 1974). The notable exception is the situation when non-stationary dependent and explanatory variables form a cointegration set (Engle/Granger 1987). While it is a general agreement that macro-economic variables such as income levels and employment rates are integrated of the first order (i.e. I(1) variables), there still is limited evidence on the time series properties of the migration flow- and corresponding migrant stock variables. Hatton (1995) provides in his analysis of the UK-US migration episode between 1870 and 1913 evidence that all variables in equation are I(1), but it is unclear whether this is supported also by other data sets. Particularly puzzling is the fact that both the migration flow and the migration stock variables are included in equation (2). Since migration flows can be conceived as (almost) the first difference of migration stocks, they can hardly be I(1) variables if migration stocks are supposed to be I(1) variables as well. In contrast, under the theoretical considerations of the stock model, it can be expected that migration stocks are I(1) variables, while the (net) migration rate can be approximated by an I(0) process. This is tested below.

Data

The empirical analysis is based on two samples. The first sample comprises the migration data from 18 European source countries to Germany from 1967 to 2001.[5] This sample covers all European source countries with the exception of the countries of the former COMECON and Albania, and the successor states of the Yugoslavia. The first group of countries has been excluded since the 'iron curtain' has effectively restricted migration from there for most of the sample period, and the second group since the (civil) wars in the former Yugoslavia hinders a meaningful analysis of the economic forces which drive migration. Germany has been chosen as the destination country for two reasons. First, it is, at some 40 % of the foreign residents in the EU, the largest destination for migrants in the Community. Second, Germany is one of the few European countries which report both migration stock and flow data by country of origin since 1967, which allows to apply the tools of modern time-series econometrics.

Both the migration stock and flow data are characterised by a visible structural break in 1973, i.e. the year of the first oil-price shock. The bilateral guest-worker agreements between Germany and a number of important source countries have been dispensed at the same year. However, the same structural break is also observed for migration from the EU Member States which have not been affected by the change in the institutional conditions. Since this structural break and changes in the institutional conditions can affect the unit-root and cointegration tests, we perform the tests also on a second sample where the migration data is not affected by structural breaks. This sample covers the period from 1973 to 2001 and includes the founding members of the Community (Belgium, France, Itlay, Luxembourg, Netherlands), the three countries of the first enlargement round (Denmark, Ireland, UK), and Austria and Switzerland in 1973. For the EU-members free movement was granted for the total sample period, in case of the two German speaking countries Austria and Switzerland bilateral agreements granted *de facto* free movement during the sample period.

The data on migration stocks and flows come from the Federal Statistical Office ('Statistisches Bundesamt') in Germany. For the stock of migrants, the foreign residents as reported by the Central Register of Foreigners ('Ausländerzentralregister') are used

[5] This sample comprises the 14 other members of the 'old' EU, Iceland, Norway, Switzerland and Turkey.

as a variable. This data is available from 1967 to 2001. The stock of foreign residents is reported on December 31 (in some early years on September 30). The number of foreign residents is slightly overstated by the Central Register of Foreigners, since return migration is not completely registered by the municipalities. Consequently, the figures for the stock of foreign residents has been revised two times in the wake of the population censuses in 1972 and 1987. In the econometric analysis, dummy variables are used in order to control for these breaks. Moreover, after German unification, complete figures for Western Germany are no longer available. Since the number of foreigners in Eastern Germany has been fairly low, this does not affect the total figures much. The data on migration flows stem again from the Central Register of Foreigners. The migration stock and flow variables are calculated as shares of the corresponding home population.

Population figures are depicted from the World Development Indicators 2003 (World Bank 2004). As a proxy for wages and other incomes, we employ the per capita GDP measured in purchasing power parities. Before 1995, historical time-series from Angus Maddison (1995) are used, which have been extrapolated by the real growth rates of the PPP-GDP per capita from the Main Economic Indicators of the OECD (OECD 2003). The employment rate is defined as one minus the unemployment rate. Unemployment rates have been taken again from the OECD Main Economic Indicators, and, if not available, complemented by data from national statistical offices. The ILO definition has been used for all unemployment rates.

The error correction model which is estimated in the following section includes three dummy variables for the institutional conditions to migrate: (i) $GUEST_{it}$, which has a value of one if a guestworker agreement between Germany and the respective country exists, and zero otherwise, (ii) $FREE_{it}$, which is one, if migration from this country is subject to free movement in the EU or the EEA, and zero otherwise, and (iii) $DICT_{it}$, which is one, if the political system of the respective country is characterised by dictatorship and zero otherwise. Details on the institutional variables are available from the authors upon request.

The descriptive statistics for both samples is presented in annex Table A1.

Testing for unit-roots and cointegration

In the first step of the empirical analysis, the variables are tested for unit-roots for making inference on their order of integration. To this end, the panel unit-root test suggested by Im, Pesaran and Shin (2003) (IPS-test) is applied to the variables of the alternative migration models. Note that panel unit root tests have a much higher power than the standard Augmented Dickey Fuller (ADF)-test for individual time-series, particularly if the root is close to one. The auxiliary regressions include either an intercept only or an intercept together with a linear deterministic time trend. Since trending behaviour cannot be ruled out *a priori*, we report for all variables the results for both sets of deterministic components. The lag-length has been chosen by the modified Schwarz criterion.

Table 2 reports the results of the IPS-test for the variables of both the stock and the flow migration model.[6] As expected, the null hypothesis that the macroeconomic variables, i.e. the relative income ratio and the employment rates, follow I(1) processes, cannot be rejected.[7] Moreover, the null of an I(1) process cannot be rejected for the migrant stock variable as well. In case of the net migration rate, it is striking that the hypothesis of a unit root is rejected at the 1 %-significance level both in the ADF-regression with an intercept only and with an intercept and a deterministic trend. For both samples, i.e. for the larger country sample which might be affected by structural breaks, and for the smaller sample of the ten source countries for which free movement was granted or *de facto* granted in the period 1973–2001, we obtain similar results.

Thus, the assumption of the standard migration model that net migration flows and macroeconomic variables such as GDP per capita levels or employment rates are integrated of the same order is not supported by the data set employed here. As a consequence, the flow model in equation (2) is unbalanced as the net migration rate, which has been found to be an I(0) variable, is being explained by non-stationary I(1) variables.

In order to reconcile the features of the data with the theoretical considerations, the long-run migration function of the migration stock model as specified in equation (3) is employed for the further analysis. According to the unit-root test results, all the variables of equation (3) seem to be I(1), such that

[6] The results of the ADF-tests for the individual time series are available from the authors upon request.

[7] For the German employment rate the individual ADF-test statistic is in the regression with intercept −2.028 (at 2 lags), which yields a p-value of 0.27, and in the regression with intercept and deterministic trend −1.509 (without lags), which gives a p-value of 0.80.

Table 2
Panel unit-root test results

	intercept		intercept and trend	
variable	W-statistic	p-value	W-statistic	p-value
	18 source countries, 1967–2001			
mst_{it}	–1.098	0.136	0.239	0.594
m_{it}	–4.895***	0.000	–2.428***	0.008
$\ln(w_{ft}/w_{it})$	4.238	1.000	1.738	0.959
$\ln(w_{it})$	–0.143	0.443	–0.809	0.209
$\ln(e_{it})$	6.202	1.000	6.202	1.000
	10 source countries, 1973–2001			
mst_{it}	–0.751	0.226	0.185	0.573
m_{it}	–5.861***	0.000	–3.198***	0.001
$\ln(w_{ft}/w_{it})$	–0.339	0.367	1.767	0.961
$\ln(w_{it})$	4.529	1.000	–0.310	0.378
$\ln(e_{it})$	–0.293	0.385	5.854	1.000

*Notes: (i) The symbols ***,**,* indicate levels of 1%, 5 % and 10% respectively. (ii) The W-statistic of the Im-Pesaran-Shin test has a standard normal distribution asymptotically. (iii) The lag-length of the individual unit-root tests was chosen by the modified Schwarz criterion.*

they can hypothetically form a cointegration set. Under the assumption of cointegration, the remainder term ε_{it} is assumed to be an I(0) variable.

Table 3 presents two cointegration tests. The first test comprises the results of the two-step Engle-Granger cointegration procedure performed for the variables of every country. The second test comprises the panel cointegration group *t*-test statistic of Pedroni (1999) which aggregates the test statistics obtained in the first place for every country in the panel. In the larger sample, the null hypothesis of no cointegration is rejected in four out of the eighteen individual cointegration regressions, and in the smaller sample in three out of ten regressions. However, these tests have rather low power, particularly in case of the relatively short time-dimension of the data at hand. The application of the more powerful panel cointegration test leads to the conclusion that we can reject the null hypothesis of no cointegration at the 10 % significance level in both samples.

Thus, the results of the cointegration tests suggest that the country specific variables form a cointegrated set, although the significance level is not very high. This might be attributed to the rather short time dimension of the sample. Nevertheless, based on the results of the unit-root and cointegration tests, the model in equation (3) can be estimated in order to make inference on the parameter values of the cointegrating relations.

3 Comparing alternative estimation procedures

There are different procedures to estimate both the long-run cointegration relationship and the short-run dynamics. If the variables form a cointegrated set, the cointegrating vector can be consistently estimated in a static regression, which completely omits the dynamics of the model (Engle/Granger 1987). Although the super-consistency result (Stock 1987) indicates that convergence is rather fast, the distribution of the least squares estimator and the associated *t*-statistic is not normal in finite samples (see e.g. Patterson 2000, for details). Monte Carlo-evidence suggests that the estimation bias of the cointegrating parameter is smaller in dynamic than in static models (Banerjee *et al*. 1986). The empirical equation is therefore here specified in form of an

Table 3
Cointegration test results

	sample: 18 source countries, 1967–2001		sample: 10 source countries, 1973–2001	
	test-statistic	lags	test-statistic	lags
AUS	–5.359***	0	–4.352	1
BEL	–3.084	4	–3.166	2
DK	–2.653	1	–2.765	2
ESP	–5.629***	3	–	
FIN	–3.853*	1	–	
FRA	–2.801	1	–3.019	2
GRE	–3.380	4	–	
ICE	–3.452	3	–	
IRE	–3.211	2	–3.517	2
ITA	–2.640	2	–2.823	2
LX	–4.302**	4	–4.203**	1
NET	–3.543	0	–4.400**	2
NOR	–2.750	0	–	
POR	–3.039	4	–	
SWE	–2.430	0	–	
SWI	–3.721	0	–3.540	1
TK	–2.737	1	–	
UK	–3.282	2	–2.974	2
group *t*-statistics	**–1.588***		**–1.348***	

*Notes: (i) The symbols ***,**,* indicate levels of significance of 1%, 5%, and 10% respectively. (ii) In the regressions with intercept, the critical values for the rejection of the null of a unit-root are -4.73, -4.11 and -3.83 at the 1%, 5%, and 10% significance level respectively. (iii) In the regressions with intercept and deterministic trend, the critical values for the rejection of the null of a unit-root are are –4.65, –4.16 and –3.84 at the 1%, 5%, and 10% significance level respectively. (iv) The group t-statistics have the asymptotic standard normal distribution. The critical values for the null hypothesis of a unit-root are –2.63, –1.64 and –1.28 at the 1%, 5% and 10% significance level respectively.*

error correction model (ECM), which allows to estimate both the long-term cointegrating vector and the short-run dynamics. Note that the ECM is a very flexible functional form and imposes few restrictions on the adjustment process.

Specifically, the estimation model has the form

$$\begin{aligned}\Delta mst_{it} = {} & \beta_{i1} mst_{i,t-1} + \beta_{i2} \ln(w_f/w_i)_{t-1} + \beta_{i3} \ln(w_i)_{t-1} \\ & + \beta_{i4} \ln(e_f)_{t-1} + \beta_{i5} \ln(e_i)_{t-1} + \beta_{i6} \Delta \ln(w_f/w_i)_t \\ & + \beta_{i7} \Delta \ln(w_i)_t + \beta_{i8} \Delta \ln(e_f)_t + \beta_{i9} \Delta \ln(e_i)_t \\ & + \sum_{j=0}^{p} \delta_{ij} \Delta mst_{i,t-1-j} + \boldsymbol{\eta}_i' \mathbf{z}_{it} + \mu_i + \varepsilon_{it}, \qquad (4)\end{aligned}$$

where $\mathbf{z}_{it}$ is a vector of institutional variables, $\boldsymbol{\eta}_i$ is the corresponding vector of coefficients, and Δ is the first difference operator. The parameter $-\beta_{i1}$ determines the speed of adjustment and the long-term coefficients of the cointegrating relationship are given by $-\beta_{in}/\beta_{i1}$, where $n = 2, 3 \ldots 5$. We consider three dummy variables here which should capture different institutional conditions for migration: guestworker agreements between the source country and Germany, free movement between the source country and Germany, and dictatorship in the source country. The first two variables should capture reduced legal and administrative barriers for migration, the last variable a political 'push' factor in the source country. If possible, time-invariant variables such as distance and dummy variables for geographical proximity and common language are considered as well.

A fundamental question for the estimation of the model in equation (4) is whether the data should be pooled or not, i.e. whether the country specific parameters are restricted to be uniform ($\beta_i = \beta, \forall \beta_i$).

Pooling can produce inconsistent and potentially misleading estimates of the parameter values unless the slope coefficients are identical (Pesaran/Smith 1995). However, there are few examples in the econometric literature where the homogeneity assumption of pooled models cannot be called into question. Several alternatives to the pooling of the data can be considered. The regressions can be estimated individually and the means of the estimated coefficients calculated. This 'Mean Group' estimator produces consistent results if the group dimension of the panel tends to infinity (Pesaran/Smith 1995) – which is however not the case in the sample at hand. Another alternative is the 'Pooled Mean Group' estimator, which constrains the long-term coefficients to be the same but allows for heterogeneous short-run coefficients. This estimator is an intermediate case – it imposes less restrictions on the adjustment process, but the same restrictions on the long-term coefficients as standard panel models. If the variables of the model form a cointegrated set, similar assumptions on the convergence of the estimated parameters as in individual regressions apply for the 'Mean Group' and the 'Pooled Mean Group' estimators (Pesaran *et al.* 1997).

Although the theoretical arguments against the homogeneity assumption of pooled estimators are appealing, there exists ample evidence that the forecasting performance of traditional panel estimators such as fixed effects and pooled OLS estimators is superior relative to heterogeneous estimators in many empirical applications (Baltagi *et al*. 2002; Baltagi *et al*. 2000; Baltagi/Griffin 1997). The reason for this finding is that individual regressions can yield highly unstable results if data sets have a limited time-dimension.

Against this background, different pooled and heterogeneous estimation procedures are applied in this paper. Six groups of estimators are considered:

- Pooled OLS (POLS) estimators;
- Fixed effects (FE) estimators;
- Random effects (RE) estimators;
- Generalized Methods of Moments (GMM) estimator;
- Pooled Mean Group (PMG) estimator;
- Mean Group (MG) estimator;
- Individual OLS (IOLS).

The POLS estimator, which imposes homogenous intercept and slope coefficients, restricts not only the slope parameters, but also the intercept to be uniform across countries. It can be applied both with and without time-invariant variables, the first estimator is labelled here POLS and the second POLS(TINV). In the first case individual (country) specific effects are completely ignored, in the latter case they are only considered to the extent they are captured by the time-invariant variables of the model.

The second group of estimators treats country specific effects as fixed. The fixed effect estimator is based on the within transformation of the data that wipes out all time invariant variables. We employ three types of the fixed effects estimator here: ones that allow only for the homogenous disturbances (FE), for group-wise heteroscedasticity FE(HET) and both for group-wise heteroscedasticity and cross-sectional correlation FE(HET + COR) in the disturbances.

The third group of estimators treat the country-specific effects as random and therefore also employs variation between the cross-sections. Depending on the way the optimal weights are attached to the within and between variation, one can distinguish between the random effects estimator of Wallace/Hussein (1969), RE(WALHUS), and the iterative GLS estimator, RE(MLE), which is equivalent to the Maximum Likelihood Estimator. Due to insufficient degrees of freedom in the between dimension, the standard Swamy/Arora (1972) random effects estimator is not employed here.

The GMM estimator addresses the fact that simultaneous equation bias caused by the presence of the lagged dependent variable can affect the estimation of dynamic panel models (Nickell 1981, Kiviet 1995). Although the simultaneous equation bias disappears with the time dimension of the panel, it can still be relevant for the size of our panel with slightly more than thirty observations over time (Judson/Owen 1999). Therefore the GMM-estimator by Arellano/Bover (1995) (GMM-SYS) is applied here, which addresses the simultaneous bias by using the appropriate set of instruments. The Arellano/Bover (1995) estimator employs both the first differences and levels equations. Blundell/Bond (1998) report Monte Carlo evidence and empirical results which indicate that the Arellano/Bond (1995) system estimator achieves substantial efficiency improvements relative to the Arellano/Bond (1991) first difference estimator. However, gains from unbiased estimation through GMM procedures might be offset by losses in efficiency if the time dimension is large relative to the time dimension of the data set (Baltagi *et al*. 2000).

Table 4

Regression results: long-run semi-elasticities

	POLS	POLS-TINV	FE	FE (HET)	FE (HET+COR)	RE (WALHUS)	RE (MLE)	GMM (SYS)	PMG	MG
	t-statistics in parentheses									
$\ln(w_f/w_h)_{t-1}$	6.146*	2.019	0.584**	0.330*	0.588***	5.812	0.947**	3.263***	0.451***	0.280
	(1.67)	(1.22)	(2.00)	(1.89)	(5.22)	(1.48)	(2.44)	(4.53)	(3.00)	(0.13)
$\ln(w_h)_{t-1}$	4.190	1.567	0.698***	0.444***	0.694***	3.910	0.775***	0.885	0.441***	0.440
	(1.23)	(1.57)	(2.86)	(3.42)	(6.34)	(1.02)	(2.78)	(1.21)	(3.00)	(0.56)
$\ln(e_f)_{t-1}$	48.860	19.133	4.897**	2.709***	4.281***	45.776*	5.546***	5.648	3.151***	3.357
	(1.08)	(1.80)*	(2.80)	(3.33)	(6.85)	(1.73)	(3.33)	(1.36)	(3.87)	(0.71)
$\ln(e_h)_{t-1}$	1.041	–0.557	–1.088*	–0.842**	–0.915***	0.137	–1.165	–1.141	–0.831***	–1.783
	(0.19)	–(0.29)	–(1.97)	–(2.19)	–(10.86)	(0.02)	–(1.48)	–(0.53)	–(7.91)	–(0.45)

*Notes: The symbols ***,**,* indicate levels of significance of 1%, 5%, and 10% respectively.*

The last three groups represent the heterogeneous estimators. As discussed before, the Pooled Mean Group (PMG) estimator imposes the restriction of common coefficients for the long-run coefficients but allows for heterogeneous short-run coefficients, while the Mean Group (MG) estimator is based on the averages of the coefficients in individual regressions. The individual OLS estimator (IOLS) allows for heterogeneous intercepts, short- and long-run coefficients.

The choice of estimators here reflects three purposes: First, pooled estimators that ignore individual specific effects are confronted with pooled estimators which include country specific effects. This has resulted in considerable differences in the long-run coefficients and, hence, in the forecasts in the migration literature. Second, GMM estimators which address the problem of simultaneous equation bias are compared to traditional panel estimators which ignore this problem. Third, estimators that allow for complete heterogeneity amongst the different cross-sections are confronted with panel estimators which rely on the fundamental assumption of homogenous slope parameters.

Table 4 presents the estimation results for the long-run semi-elasticities for the panel estimators. The estimation results for the short-run semi-elasticities are reported in Annex Table A2, and the long-run semi-elasticities of the individual OLS regressions in Annex Table A3.[8] Before discussing the estimation results for every group of estimators, it is worthwhile summarising the general observations:

First, the coefficients of the lagged migration stock full-fill the conditions for dynamic stability for all panel and all individual OLS estimates with the exception of Ireland.

Second, for a number of the estimators (POLS, POLS-TINV, RE(WALHUS)) the coefficient for the lagged migration stock variable is very close to zero, implying a very high degree of persistence of the underlying time series. This fact also results in extremely high values of the long-run coefficient estimates – their values are around ten-times as high as those of the fixed-effects estimates. This corresponds to the extremely high estimates which are obtained by forecasts of the migration potential from Central and Eastern Europe based on pooled OLS estimators.

Third, the estimated coefficients for the foreign-to-home wage ratio, w_f/w_i, the home wage, w_i, and the German employment rate, e_f, have the expected positive sign and appear significant for almost all panel data estimators. However, for the home employment variable, e_h, the coefficient has in some regressions an unexpected positive sign and appears frequently insignificant.

Fourth, the individual OLS regressions yield very heterogeneous coefficient estimates. Consequently, the coefficients of the Mean Group estimator appear insignificant for all variables except the lagged migration stock.

[8] The estimation results for the short-run semi-elasticities of the individual OLS regressions are available from the authors upon request.

Finally, among the institutional variables guestworker recruitment and dictatorship have the expected signs and appear significant in most panel regressions. However, the coefficient of the free movement dummy is frequently insignificant and has sometimes an unexpected sign. In principle, the impact of guestworker recruitment agreements appears to be much larger than the impact of the free movement with the EU.

The first group of estimators include the pooled OLS with and without the time invariant variables. Since these estimators ignore country-specific effects (apart from those captured by the time-invariant variables), the results are certainly biased if the country-specific effects are correlated with the explanatory variables. Note that a correlation between the lagged dependent variable and country specific effects can explain why the coefficient of the lagged dependent variable tends to zero (see e.g. Baltagi *et al.* 2000).

The second group comprises the fixed effects estimators: FE, FE(HET), and FE(HET + COR). As the regression diagnostics demonstrates, a standard *F*-test clearly rejects the null hypothesis of a uniform intercept at the 1 %-significance level. Moreover, the Likelihood Ratio test results indicate that both, group-wise heteroscedasticity and cross-sectional correlation in the disturbances is present in the sample. As seen, the estimates based on the within transformation yield much higher negative values for the lagged migration stock indicating a lower persistence over time and faster speed of adjustment to the long-run equilibrium.

The random effects estimators yield mixed results. On the one hand, the RE(WALHUS) estimates are very close to those of the pooled OLS estimators. This can be explained by the fact that the optimal weighting for the RE(WALHUS) is based on the OLS residuals (see Doornik *et al*. 2002). On the other hand, the RE(MLE) estimates are similar to those of the fixed effects. This can be traced back to the fact that the importance of the within variation in the GLS optimal weighting scheme increases with the growing time dimension (see Baltagi 1995).

The GMM-estimator yields similar results for the lagged migration stock as the fixed effects estimators, but much higher values for the coefficients of the explanatory variables such as the wage ratio. The results of the PMG-estimator, which imposes the same restriction on the long-run coefficients as the panel estimators, but allows the parameters of the short-run variables to differ across countries, yields results which are very similar to the fixed effects estimates.

Finally, the individual OLS regressions produce on average much larger coefficients for the lagged migration stock variable relative to the panel-estimators, as expected by econometric theory. The MG estimator yields therefore an estimate for the parameter of the lagged migration stock of −0.255, which is the largest among the estimators considered. However, the results for the explanatory variables are very heterogeneous and in most cases insignificant. Given the rather small time dimension of the sample, the regression results are highly unstable.

Forecasting performance

For evaluation of the out-of-sample forecasting performance of the different models, the Root Mean Squared Forecast Error (RMSFE) is calculated in Table 5. The out-of-sample forecasting performance is compared for two time periods: the fifth and the tenth year ahead. For this purpose the estimated coefficient values on the samples 1969–1996, and 1969–1991, respectively, have been employed.

The ranking of the estimators based on both measures of the forecast accuracy varies slightly with the forecasting horizon. The fixed effects estimators dominates in both cases all other estimators, but the FE(HET + COR) is replaced on the first place by the FE(HET) estimator in the longer forecasting period. The RE(MLE) estimator, which places a high weight on the within variation in the data and yields therefore results which are close to the fixed effects estimators, performs relatively well in the shorter forecasting period, but the forecasting accuracy declines substantially in the longer forecasting period. The other random effects estimator, the RE(WALHUS) estimator, which addresses a much higher weight on the between variation, performs poorly in both time horizons. The PMG estimator, which yields similar estimates of the long-run coefficients as the fixed effects estimators, has a mediocre performance in both time periods, with a forecasting error which is twice as high as that of the fixed effects estimators. The GMM(SYS) estimator performs similar as the PMG estimator, indicating that the gains from unbiased estimation do not offset the losses in efficiency for the data set employed here. The pooled OLS estimators, which are popular in the literature, perform with a forecasting error which is around three time higher than that of the fixed effects estimators at the lower end of the estimators. Interestingly enough, the individual OLS estimator, performs poorly in the 5-year time period but improves the forecasting accuracy in the 10-year time period substantially.

Nevertheless, the forecasting error is still substantially higher than that of the fixed effects estimators.

Table 5

Forecasting performance of different estimators

5 years ahead			10 years ahead		
rank		RMSFE	rank		RMSFE
1	FE(HET+COR)	0.0860	1	FE(HET)	0.1061
2	FE(HET)	0.0925	2	FE(HET+COR)	0.1333
3	FE	0.0987	3	FE	0.1377
4	RE(MLE)	0.1495	4	IOLS	0.1640
5	PMG	0.1872	5	PMG	0.2240
6	GMM(SYS)	0.2112	6	GMM(SYS)	0.2390
7	POLS(TINV)	0.2115	7	RE(MLE)	0.2626
8	RE(WALHUS)	0.2129	8	POLS	0.2987
9	POLS	0.2137	9	RE(WALHUS)	0.3028
10	IOLS	0.2823	10	POLS(TINV)	0.3100
11	MG	0.8955	11	MG	0.8720

Note: RMSFE: Root mean squared forecast error.

Finally, the mean group estimator performs extremely poorly, which can be traced back to the fact that the results from the individual regressions are very heterogeneous and unstable.

These results are comparable to those of a number of studies including Baltagi/Griffin (1997), Baltagi *et al.* (2000) and Baltagi *et al*. (2002), where homogenous estimators that allow for individual specific effects (e.g. fixed effects estimators) outperform heterogeneous estimators such as individual OLS and Mean Group estimators. This can be explained by the fact that in cases where the results of individual estimates are relatively unstable over time and groups (i.e. countries) standard panel estimators tend to dominate individual estimators.

4 Potential migration in an enlarged EU

In this section we examine the migration potential both for countries within the sample as well as out of the sample, i.e. the accession countries from Central and Eastern Europe. Within sample, the long-run coefficients of the regressions allow to calculate the long-run migration potential for the countries considered by the estimates. The countries included in the sample comprise the members of the EEA, for which the free movement of labour is already granted, and Turkey.

Table 6 reports the long-run migration potential as predicted by the different estimators for the countries within sample, based on the values for the explanatory variables in the last year of the estimation period (2001). The fixed effects estimators, which have proved to have the best forecasting performance, suggest that actual migration stocks have achieved already their long-run values in case of most source countries.

Even in the case of Turkey, most estimators predict that the actual migration stock is at currently 2.85 per cent of the home population already relatively close to its long-run equilibrium level. The main exceptions are the individual OLS and the GMM estimators, which predict a long-run migration potential of 4.5 and 3.85 per cent of the home population, respectively. However, the actual migration potential is hard to predict, since all these estimates reflect the current legal and administrative restrictions to immigration from Turkey. A reliable estimate of the long-run migration potential for Turkey under the conditions of free movement is hard to obtain from our estimates. The free movement dummy is identified only by the integration of relatively rich countries into the EU. As a consequence, very low coefficients have been obtained for the free movement dummy in all regressions. This might be traced back to the fact that migration stocks of the countries in the sample have been already very close to their equilibrium levels when free movement was granted, such that the liberalisation has not much affected migration behaviour.

Table 6

Forecasts of the long-run migration potential

	FE (HET+COR)	FE (HET)	FE	RE (MLE)	PMG	GMM (SYS)	POLS-TINV	RE (WALHUS)	POLS	IOLS	MG	actual 2001
AUS	2.46	2.42	2.45	0.95	2.17	0.43	1.99	1.38	1.44	2.35	0.97	2.33
BEL	0.30	0.26	0.27	0.95	0.24	0.41	1.51	1.37	1.43	0.23	0.97	0.23
DK	0.40	0.36	0.38	0.92	0.41	0.32	1.42	1.32	1.39	0.45	0.95	0.40
ESP	0.39	0.35	0.38	1.08	0.36	1.14	1.65	1.88	1.89	0.28	1.06	0.32
FIN	0.41	0.34	0.40	0.99	0.32	0.53	1.52	1.44	1.47	0.31	1.02	0.31
FRA	0.27	0.23	0.26	0.97	0.21	0.32	1.47	1.28	1.32	0.17	1.02	0.19
GRE	3.42	3.41	3.42	1.08	3.41	1.60	2.02	2.28	2.33	3.33	0.97	3.43
ICE	0.57	0.51	0.56	0.88	0.73	0.32	1.39	1.36	1.46	0.55	0.88	0.53
IRE	0.36	0.34	0.32	0.89	0.35	0.10	1.35	1.15	1.23	0.58	0.94	0.41
ITA	1.16	1.12	1.15	1.00	1.04	0.65	1.69	1.53	1.57	0.97	1.02	1.07
LX	1.44	1.42	1.42	0.82	1.46	–0.66	1.29	0.57	0.65	1.38	0.96	1.41
NET	0.83	0.78	0.81	0.90	0.68	0.43	1.49	1.44	1.54	0.70	0.89	0.70
NOR	0.29	0.22	0.28	0.90	0.22	0.20	1.37	1.24	1.32	0.17	0.93	0.17
POR	1.14	1.13	1.13	0.99	1.17	1.34	1.77	2.15	2.25	1.41	0.86	1.32
SWE	0.31	0.24	0.31	0.93	0.27	0.43	1.42	1.41	1.48	0.21	0.94	0.22
SWI	0.60	0.55	0.59	0.81	0.59	0.08	1.24	0.63	0.70	0.53	0.92	0.53
TK	2.92	3.04	2.89	1.14	3.28	3.85	2.12	3.58	3.69	4.51	0.77	2.84
UK	0.26	0.22	0.25	0.94	0.21	0.60	1.48	1.54	1.62	0.22	0.93	0.19

Note: All forecasts are based on the 2001 values of the explanatory variables. See text for assumptions.

Thus, without a major change in the institutional migration conditions, the migration potential from the European source countries seem to be already largely exhausted for Germany according to our estimates. Eastern Enlargement of the EU is such a major institutional change. As discussed before, the Central and Eastern European countries have not been included in the sample due to the iron curtain, which effectively prevented emigration for most of the time period under consideration. Thus, on basis of our estimates, we have to rely on an out-of-country projection. This involves further problems, especially in case of the fixed effects estimators. Fixed effects are by definition country-specific variables and cover both observable and non-observable factors. We follow here the procedure by Fertig (2001) and Boeri/Brücker (2001) and explain the fixed effects by time-invariant variables. More specifically, we use distance and distance squared, a dummy variable for geographic proximity (*ADJACENT*), a dummy variable for a location in the eastern part of Europe (*EAST*), and a dummy variable for common language as explanatory variables. The adjusted R^2 statistics suggest that almost 90 percent of the variance in the fixed effects is explained by these variables (see Annex Table A4).

Moreover, the migration scenario is based on the following assumptions for the explanatory variables: the per capita GDP in Germany grows at an annual rate of 2 %; the per capita GDP of the accession countries converges to that of Germany and the EU-15 at an annual rate of 2 %; the (un-)employment rates in Germany and the accession countries remain constant. Note that the convergence rate of 2 % p.a. for the per capita GDP corresponds to that found by Barro/Sala-i-Martin (1991, 1995) and in many other studies for the EU and other European market economies. The growth rates of the accession countries fit pretty well into this picture since the end of the transition recession.

Under these assumptions, the FE(HET+COR) estimator predicts an initial immigration of around 225,000 persons from all ten accession countries from Central and Eastern Europe (the eight new Member

Table 7

Projection of the migration potential from the accession countries (ACs) to Germany, 2004–2030

	2004	2005	2006	2007	2008	2009	2010	2015	2020	2025	2030
	net migration (1,000 persons)										
AC-10	225	258	240	203	164	128	98	29	16	13	10
AC-8	156	169	155	132	108	86	68	24	13	10	7
AC-2	70	89	84	71	56	42	31	6	3	3	3
	foreign population (1,000 persons)										
AC-10	824	1082	1322	1525	1689	1817	1915	2159	2258	2327	2384
AC-8	628	797	952	1084	1192	1278	1345	1527	1608	1664	1705
AC-2	196	285	370	441	497	539	570	632	649	663	679

Notes: AC-10: Bulgaria, Czech Republic, Estonia, Hungaria, Latvia, Lithuania, Poland, Romania, Slovakia, Slovenia. - AC-8: AC-10 without Bulgaria and Romania. - AC-2: Bulgaria, Romania.-- See text for assumptions of the projection.

States and Bulgaria and Romania) if the free movement would have been hypothetically introduced for all countries in 2004. The long-run migration potential is achieved at around 2.4 million persons around 25 years later (Table 7). According to the FE(HET) estimator, which is not reported here, the initial immigration would be at around 185,000 persons and the long-run migration potential at around 2.2 million persons.

5 Conclusion and policy-implications

Forecasts of the migration potential from Central and Eastern Europe differ largely. In this paper, several methodological and empirical problems in estimating macro migrations models and forecasting migration potentials have been analysed. The methodological aspects of our analysis can be summarised as follows: first, our results suggest that the standard migration model, which explains migration flows by income and employment variables and (lagged) migration stocks, is at least in case of our data set not properly specified, since the migration flow variable seem to be stationary, while the explanatory variables seem to be I(1). As a consequence, the conditions for a long-run equilibrium relationship are violated. Instead our data set suggests that there seem to be cointegration relationship between migration stocks and the explanatory variables of the standard migration model. However, one caveat applies to our findings: we can reject the null hypothesis of no cointegration only at the 10 %-significance level. Probably, this can be traced back to the rather short time dimension of the panel.

Second, by testing a large set of panel and heterogeneous estimators, we find that the forecasting accuracy of standard fixed effects estimators outperforms (i) pooled OLS estimators and random effects estimators, which are widely applied in the migration context, (ii) GMM-system estimators, which try to avoid the simultaneous equation bias of dynamic panel models with finite time dimension, (iii) heterogeneous estimators, which relax by one way or another the fundamental homogeneity assumption of panel estimation procedures.

These technical findings have a number of policy consequences. First, the flow model suggests that migration ceases not before (expected) income levels between the host and the source country have converged to a certain threshold level, which is determined by the costs of migration. In case of persistent differences in (expected) income levels, either the total population will eventually migrate or migration will not happen at all from the beginning. In contrast, the stock model predicts that migration ceases when the benefits from migration equals its costs for the marginal migrant, such that a long-run equilibrium between migration stocks and the income variables emerges. This helps to explain why the introduction of the free movement in case of the EU's Southern enlargement has not involved any significant increase in migration: in this case, the equilibrium stock of migrants has been already achieved before the free movement was introduced, such that introducing the free movement had no effect.

Second, the comparison of different panel and heterogeneous estimators has some policy consequences as well. The pooled OLS estimators implicitly assume that all factors which have a persistent impact on migration, that is inter alia distance, language and culture, are the same for all source countries. Statistical tests clearly reject this assumption, and, not surprisingly, the forecasting performance of these estimators is weak. More surprising is our finding that the heterogeneous estimators are dominated by fixed effects estimators, since it is reasonable to assume that slope parameters differ between countries as well. However, the individual regression results have turned out to be rather unstable. Thus, at least in our sample of European source countries, the restriction that the slope parameters are identical increases the forecasting accuracy. Note that similar results have been obtained also in other contexts.

These findings have implications for the scale of the migration potential from the accession countries: Extremely high estimates of the migration potential, which predict on basis of the pooled OLS estimator a long-run migration potential for Germany alone of 7.2 % of the population in the source countries (Flaig 2001; Sinn *et al*. 2001), can be laid to rest, since the forecasting performance of these estimators has proved to be poor and the assumption of a common intercept is clearly rejected by the specification tests. On basis of the fixed effects estimators, which have the best forecasting performance in our sample, the long-run migration stock from the accession countries amounts to 2.2 – 2.4 million persons for Germany, which implies a net immigration of another 1.6 to 1.8 million persons at a present population of 600,000 from this region. The long-run migration potential will be realised in a period of around 20 years.

Moreover, our quantitative findings suggest that migration stocks within the EU-15 are already at or very close to their equilibrium levels. This finding is confirmed by the fact that international migration within the EU-15 is low. Thus, we do not have to expect that migration will accelerate in this area in the future. Even in the case of Turkey, fears of a mass-migration wave seem to be ill-founded. Most estimators suggest that the migration potential is already relatively close to its equilibrium. The major exception is the individual OLS estimator, which predicts that the long-run equilibrium stock amounts to 4.5 % of the Turkish population, while at present 2.8 % of the Turkish population reside in Germany.

Needless to say, all these estimates are based on a number of artificial assumptions and can provide no more than a hint to the actual magnitudes involved. The heterogeneity in the migration behaviour across countries increases in particular the uncertainty on the migration potential from countries which are not included in the estimation sample such as the accession countries from Central and Eastern Europe. Moreover, the different application of transitional periods across the EU members might result in the diversion of migration flows, which in turn might yield a reduced migration potential when the transitional periods expire in Germany. Thus, all forecasts of future migration flows and stocks from the accession countries have to be treated with great caution.

References

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

Alecke, B./Huber, P./Untiedt, G. (2001): What a difference a constant makes: how predictable are international migration flows? OECD (Ed.): Migration Policies and EU Enlargement. The Case of Central and Eastern Europe. Paris, pp. 63 – 78.

Arellano, M./Bond, S. (1991): Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. In: Review of Economic Studies, Vol. 58, pp. 277 – 97.

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

Baltagi, B. H./Bresson, G./Pirotte, A. (2002): Comparison of Forecast for Homogeneous, Heterogeneous and Shrinkage Estimators. Some Empirical Evidence for US Electricity and Natural-Gas Consumption. In: Economics Letters, Vol. 76, pp. 375 – 82.

Baltagi, B. H./Griffin, J. M./Xiong, W. (2000): To Pool or of Not to Pool: Homogeneous versus Heterogeneous Estimators Applied to Cigarette Demand In: Review of Economics and Statistics, Vol. 82, No. 1, pp. 117 – 126.

Baltagi, B. H./Griffin, J. M. (1997): Pooled Estimators vs. Their Heterogeneous Counterparts in the Context of Dynamic Demand for Gasoline. In: Journal of Econometrics, Vol. 77, pp. 303 – 27.

Baltagi, B. H. (1995): Econometric Analysis of Panel Data. Chichester: John Wiley and Sons.

Banerjee, A./Dolado, J. J./Henry, D. F./Smith, G. W. (1986): Exploring EquilibriumEquilbrium Relationships in Econometrics through Static Models: Some Monte Carlo Evidence. In: Oxford Bulletin of Economics and Statistics, Vol. 48(3), pp. 253 – 77.

Barro, R./Sala-i-Martin, X. (1991): Convergence across States and Regions. Brookings Papers on Economic Activity, No. 1991/1, pp. 107 – 182.

Barro, R./Sala-i-Martin, X. (1995): Economic Growth, New York *et al.*: McGraw-Hill.

Bauer, T. (1995): The migration decision with uncertain costs. Münchner Wirtschaftswissenschaftliche Beiträge, 95–25, Munich: LMU.

Blundell, R. W./Bond, S. R. (1998): Initial conditions and moment restrictions in dynamic panel data models. In: Journal of Econometrics, Vol. 87, pp. 115–143.

Boeri, T./Brücker, H. et al. (2001): The Impact of Eastern Enlargement on Employment and Labour Markets in the EU Member States. Berlin/Brussels/Milano.

Brücker, H. (2001): Die Folgen der Freizügigkeit für die Ost-West-Migration. Schlussfolgerungen aus einer Zeitreihenanalyse der Migration nach Deutschland. In: Konjunkturpolitik – Applied Economics Quarterly, Supplement 52, pp. 17–54.

Brücker, H./Schröder, P. J. H. (2005): International Migration With Heterogeneous Agents: Theory and Evidence. International Centre for Economic Research (ICER) Working Paper 28.

Burda, M. C. (1995): Migration and the Option Value of Waiting. In: Economic and Social Review, Vol. 27 (1), pp. 1–19.

Daveri, F./Faini, R. (1999): Where Do Migrants Go? In: Oxford Economic Papers, Vol. 51, pp. 595–622.

Doornik, J. A./Arellano, M./Bond, S. R. (2002): Panel Data estimation DPD for Ox. Manual, available online:http://www.nuff.ox.ac.uk/users/doornik/

Dustmann, C./Casanova, M./Fertig, M./Preston, I./Schmidt, C. M. (2003): The impact of EU enlargement on migration flows. Home Office Online Report 25/03, http://www.homeoffice.gov.uk/rds/pdfs2/rdsolr2503.pdf.

Engle, R. F./Granger, C. (1987): Co-integration and error correction: interpretation, estimation, and testing. In: Econometrica, Vol. 66, pp. 251–76.

Eurostat (2003): Kaufkraftparitäten und abgeleitete Wirtschaftsindikatoren für EU, beitretende Länder, Beitrittskandidaten und EFTA. In: Statistik kurz gefasst, No. 64.

Faini, R./Venturini, A. (1995): Migration and Growth: The Experience of Southern Europe. CEPR Discussion Paper No. 964.

Faini, R./DeMelo, J./Zimmermann, K. F. (Eds.) (1999): Migration. The Controversies and the Evidence. Cambridge: Cambridge University Press.

Fassmann, H./Hintermann, C. (1997): Migrationspotential Ostmitteleuropa. Struktur and Motivation potentieller Migranten aus Polen, der Slowakei, Tschechien und Ungarn. ISR-Forschungsberichte, Institut für Stadt- und Regionalforschung, H. 22, Vienna.

Fertig, M. (2001): The economic impact of EU enlargement: assessing the migration potential. In: Empirical Economics, Vol. 26, pp. 707–720.

Fertig, M./Schmidt, C. M. (2001): Aggregate-Level Migration Studies as a Tool for Forecasting Future Migration Streams. Discussionspaper No. 324, University of Heidelberg, Economics Department.

Fields, G. S. (1978): Place to Place Migration: Some New Evidence. In: Review of Economics and Statistics, Vol. 61, pp. 21–32.

Flaig, G. (2001): Die Abschätzung der Migrationspotentiale der osteuropäischen EU-Beitrittsländer. In: Konjunkturpolitik – Applied Economics Quarterly, Supplement 52, pp. 55–76.

Granger, C. W./Newbold, J. P. (1974): Spurious regression in econometrics. In: Journal of Econometrics, Vol. 2(2), pp. 111–20.

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

Hatton, T. J. (1995): A Model of U. K. Emigration, 1870–1913. In: Review of Economics and Statistics, Vol. 77, pp. 407–415.

Hicks, J. R. (1932): The Theory of Wages. London: McMillan.

Hille, H./Straubhaar, T. (2001): The Impact of EU-Enlargement on Migration Movements and Economic Integration: Results of Recent Studies. In: OECD (Ed.): Migration Policies and EU-Enlargement. Paris, pp. 79–100.

Hönekopp, E. (2000): Auswirkungen der EU-Osterweiterung auf die Arbeitsmärkte der Mitgliedsländer der Europäischen Union. Report for the Friedrich-Ebert-Foundation, Bonn.

Im, K. S./Pesaran, M. H./Shin, Y. (2003): Unit root testing in heterogeneous panels. In: Journal of Econometrics, forthcoming.

IOM (International Organization for Migration) (1998): Migration Potential in Central and Eastern Europe. Geneva (IOM).

Judson, R. A./Owen, A. K. (1999): Estimating dynamic panel data models: a guide for macroeconomists. In: Economic Letters, Vol. 65, pp. 9–15.

Kiviet, J. (1995): On Bias, Inconsistency and Efficiency of Various Estimators in Dynamic Panel Data Models. In: Journal of Econometrics, Vol. 68, pp. 53–78.

Krieger, T. (2003): Migration trends in an enlarged EU. European Foundation for the Improvement of Working and Living Conditions, Dublin.

Layard, R./Blanchard, O./Dornbusch, R./Krugman, P. (1992): East-West Migration: The Alternatives. Boston: MIT Press.

Lundborg, P. (1991): Determinants of Migration in the Nordic Labor Market. In: Scandinavian Journal of Economics, Vol. 93, pp. 363–375.

Maddison, A. (1995): Monitoring the World Economy 1820–1992. Paris: OECD.

Massey, D. S./Goldring, L./Durand, J. (1984): Continuities in Transnational Migration: An Analysis of Nineteen Mexican Communities. In: American Journal of Sociology, Vol. 99(6), pp. 1492–1533.

Massey, D. S./Espana, F. G. (1987): The Social Process of International Migration. In: Science, Vol. 237, pp. 733–738.

Nickell, S. (1981): Biases in dynamic models with fixed effects. In: Econometrica, Vol. 49, pp. 1417–1426.

OECD (2003): Main Economic Indicators. CD-Rom, Paris: OECD.

Patterson, K. (2000): An Introduction to Applied Econometrics: A Time Series Approach, New York: Palgrave.

Pederson, P. J./Pytlikova, M./Smith, N. (2004): Selection or network effects? Migration flows into OECD countries 1990–2000. IZA Discussion Paper, IZA, Bonn.

Pedroni, P. (1999): Critical values for cointegration tests in heterogenous panels with multiple regressors. In: Oxford Bulletin of Economics and Statistics, Vol. 61, pp. 653–70.

Pesaran, M. H./Shin, Y./Smith, R. P. (1997): Pooled Mean Group Estimation of Dynamic Heterogeneous Panels. Mimeo, Trinity College, Dublin (revised: November 1998).

Pesaran, M. H./Smith, R. P. (1995): Estimating long-run relationships from dynamic heterogeneous panels. In: Journal of Econometrics, Vol. 68, pp. 79–113.

Ravenstein, E. (1889): The laws of migration. In: Journal of the Statistic Society, Vol. 59, pp. 214–301.

Sinn, H.-W./Flaig, G./Werding, M./Münz, S./Düll, N./Hoffmann, H. (2001): EU-Erweiterung und Arbeitskräftemigration. Wege zu einer schrittweisen Annäherung der Arbeitsmärkte. Munich: Ifo-Institut für Wirtschaftsforschung.

Sjaastad, L. A. (1962): The costs and returns of human migration. In: Journal of Political Economy, Vol. 70(5), pp. 80–93.

Stark, O. (1984): Rural-to-Urban Migration in Less Developed Countries: A Relative Deprivation Approach. In: Economic Development and Cultural Change, Vol. 32.

Stark, O. (1991): The Migration of Labour, Oxford and Cambridge. MA: Basil Blackwell.

Stock, J. H. (1987): Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors. In: Econometrica, Vol. 55, pp. 1035–56.

Stock, J. H. (1987): Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors. In: Econometrica, Vol. 55, pp. 1035–1056.

Straubhaar, T. (2002): Ost-West Migrationspotenzial: Wie groß ist es? In: Jahrbücher für Nationalökonomie und Statistik, Vol. 222, pp. 20–41.

Swamy, P. A. V. B./Arora, S. S. (1972): The exact finite sample properties of the estimators of coefficients in the error component regression models. In: Econometrica, Vol. 40, pp. 261–275.

Wallace, T. D./Hussein, A. (1969): The use of error components models in combining cross-section and time-series data. In: Econometrica, Vol. 37, pp. 55–72.

World Bank (2004): World Development Indicators. CD-Rom, Washington, D.C.: World Bank.

Table A1

Descriptive statistics

variable	obs.	mean	std. deviation	minimum	maximum
	18 source countries, 1967–2001				
mst_{it}	630	0.832	0.972	0.030	4.565
m_{it}	630	0.000	0.001	–0.006	0.007
$\ln(w_f/w_i)_t$	630	0.240	0.380	–0.383	1.589
$\ln(w_{it})$	630	9.428	0.438	7.710	10.347
$\ln(e_{ft})$	34	–0.055	0.030	–0.099	–0.006
$\ln(e_{it})$	630	–0.064	0.049	–0.277	0.000
$GUEST_{it}$	630	0.540	0.206	0	1
$FREE_{it}$	630	0.540	0.499	0	1
$DIKT_{it}$	630	0.056	0.229	0	1
	10 source countries, 1973–2001				
mst_{it}	290	0.682	0.652	0.068	2.389
m_{it}	290	0.000	0.000	–0.001	0.001
$\ln(w_f/w_i)_t$	290	0.077	0.193	–0.383	0.655
$\ln(w_{it})$	290	9.650	0.251	8.857	10.347
$\ln(e_{it})$	29	–0.065	0.024	–0.099	–0.010
$\ln(e_{it})$	290	–0.072	0.042	–0.191	0.000

Table A2

Regression results: short-run semi-elasticities

	POLS	POLS (TINV)	FE[1]	FE (HET)[2]	FE (HET+COR)[3]	RE (WALHUS)[4]	RE (MLE)[5]	GMM (SYS)[6]	PMG	MG
	t-statistics in parentheses					t-statistics in parentheses				
$\ln(w_f/w_h)_{t-1}$	0.058***	0.053*	0.087**	0.042*	0.084***	0.060	0.119**	0.381***	0.062***	0.071
	(2.97)	(1.72)	(2.21)	(1.92)	(5.32)	(1.51)	(2.43)	(3.32)	(3.03)	(.13)
$\ln(w_h)_{t-1}$	0.040**	0.041*	0.104***	0.056***	0.099***	0.040	0.097***	0.103	0.061***	0.112
	(2.13)	(1.96)*	(3.57)	(3.40)	(6.43)	(1.06)	(2.74)	(1.21)	(3.00)	(0.60)
$\ln(e_f)_{t-1}$	0.465**	0.500**	0.733***	0.342***	0.613***	0.472**	0.697***	0.659	0.435***	0.855
	(2.04)	(2.24)	(3.36)	(3.65)	(7.03)	(2.21)	(3.50)	(1.43)	(3.95)	(0.79)
$\ln(e_h)_{t-1}$	0.010	–0.015	–0.163*	–0.106**	–0.131***	0.001	–0.147	–0.133	–0.115***	–0.454
	(0.18)	–(0.29)	–(1.95)	–(2.31)	–(11.77)	(0.02)	–(1.53)	–(0.54)	–(8.49)	–(0.47)
$mst_{h,t-1}$	–0.010	–0.026***	–0.150***	–0.126***	–0.143***	–0.010***	–0.126***	–0.117***	–0.138***	–0.255*
	–(1.57)	–(2.86)	–(6.06)	–(8.66)	–(20.78)	–(3.05)	–(10.90)	–(4.35)	–(14.92)	–(1.65)
FREE	0.005	0.013	0.008	0.000	0.006***	0.006	0.009	0.029	0.003**	–
	(0.82)	(1.20)	(0.91)	(0.07)	(3.69)	(0.94)	(1.05)	(0.71)	(2.52)	
GUEST	0.158***	0.155***	0.098***	0.105***	0.109***	0.158***	0.110***	0.164***	0.101***	–
	(6.33)	(6.17)	(6.27)	(5.69)	(11.64)	(7.57)	(5.49)	(2.85)	(7.90)	
DIKT	–0.029	–0.021	0.062**	0.012	0.048***	–0.028*	0.048***	0.014	0.022**	–
	–(1.42)	–(0.98)	(2.01)	(0.77)	(5.91)	–(1.82)	(3.11)	(0.27)	(2.44)	
STAT-1972	–0.123***	–0.123***	–0.112***	–0.048***	–0.101***	–0.123***	–0.113***	–0.114***	–0.105***	–0.108
	–(2.93)	–(2.92)	–(2.97)	–(7.56)	–(15.86)	–(8.03)	–(8.26)	–(3.75)	–(13.22)	–(0.71)
STAT-1987	–0.074***	–0.075***	–0.083***	–0.048***	–0.082***	–0.074***	–0.082***	–0.080***	–0.082***	–0.082
	–(3.57)	–(3.63)	–(3.77)	–(7.79)	–(13.59)	–(5.01)	–(6.14)	–(2.74)	–(10.71)	–(0.84)
$\Delta\ln(w_f/w_h)_t$	0.083	0.058	0.102**	0.037	0.120	0.082	0.116	0.303	–	–0.115
	(1.59)	(1.02)	(2.33)	(0.33)	(1.11)	(0.31)	(0.49)	(0.57)	–	–(0.30)
$\Delta\ln(w_h)_t$	0.204	0.256**	0.358***	0.184	0.282**	0.205	0.348	0.293	–	0.119
	(1.55)	(2.06)	(2.99)	(1.57)	(2.61)	(0.75)	(1.42)	(0.51)	–	(0.36)
$\Delta\ln(e_f)_t$	1.134***	1.133***	0.851***	0.408*	0.548**	1.139**	0.896*	0.557	–	0.690
	(2.86)	(2.86)	(3.22)	(1.83)	(2.62)	(2.17)	(1.88)	(0.52)	–	(0.54)
$\Delta\ln(e_h)_t$	0.130	0.000	–0.225	–0.164	–0.163***	0.120	–0.163	0.507	–	–0.116
	(0.47)	(0.00)	–(0.97)	–(1.37)	–(5.70)	(0.46)	–(0.67)	(0.64)	–	–(0.07)
$\Delta mst_{h,t-1}$	0.454***	0.441***	0.411***	0.302***	0.410***	0.452***	0.412***	0.365***	–	0.167
	(5.02)	(4.88)	(4.75)	(7.89)	(19.03)	(13.60)	(13.50)	(4.98)	–	(0.92)
constant	–0.353**	–0.341*	–	–	–	–0.358	–0.818**	–0.954	–	–0.832
	–(2.04)	–(1.76)				–(0.98)	–(2.42)	–(1.19)	–	–(0.44)
dist x 1000	–	–1.085**	–	–	–	–	–	–	–	–
		–(2.28)								
$dist^2$ x 1000000	–	0.883*	–	–	–	–	–	–	–	–
		(1.98)								
ADJACENT	–	0.030*	–	–	–	–	–	–	–	–
		(1.75)								
LANGUAGE	–	0.045*	–	–	–	–	–	–	–	–
		(1.81)								
adjusted R^2	0.50	0.52	0.61	–	–	0.51	–	–	–	–
Log-Likelihood	–	–	899	1280	1661	–	1122	–	1759	–

*Notes: Notes: 1) The F(17, 543)-statistic for the null hypothesis that all intercepts are equal is 9.80**.-- 2) The $\chi^2(17)$-test statistic for the LR-Test of the heteroscedastic vs. the homoscedastic model is 761.04**.-- 3) The $\chi^2(153)$-test statistic for the LR-Test of the heterosecedastic and correlated vs. the heteroscedastic model is 762.58.-- 4) The $\chi^2(15)$-statistic for the Hausman-test of the random vs. the fixed effects model is 219.99, which is significant at the 1% level.-- 5) The $\chi^2(15)$-statistic for the Hausman-test of the random vs. the fixed effects model is 2.81.-- 6) The $\chi^2(115)$-statistic of the Sargan-Test for overidentifying restrictions is 85.8, which corresponds to a p-value of 0.981; the Arellano-Bond N(0,1) test statistic for no first-order autocorrelation is -5.57**, the test statistic for no second-order autocorrelation is 0.79.*

Table A3

Individual OLS regressions: long-run semi-elasticities

	$\ln(w_f/wi)_{t-1}$		$\ln(wi)_{t-1}$		$\ln(e_f)_{t-1}$		$\ln(e_i)_{t-1}$	
	coeff.	t-stat.	coeff.	t-stat.	coeff.	t-stat.	coeff.	t-stat.
AUS	1.78	0.53	0.93	1.46	0.85	0.25	3.69	0.77
BEL	0.34**	2.57	0.17***	5.73	–0.22	–0.73	0.12	0.84
DK	–0.11	–0.29	0.82***	4.92	2.27**	2.00	–0.45	–0.91
ESP	–0.30	–0.39	–0.22	–1.02	2.53**	2.09	–0.23	–0.47
FIN	–0.14	–0.48	0.14	1.45	0.87**	2.10	–0.69	–2.12
FRA	–0.24	–0.80	0.07	1.02	1.86*	1.82	–1.68	–1.78
GRE	2.85	0.96	0.44	0.31	7.85	0.90	–0.30	–0.04
ICE	1.12***	3.01	0.63***	3.81	0.85	1.22	–3.01***	–2.60
IRE	–6.87	–0.75	–4.67	–0.69	–25.02	–0.72	2.51	0.46
ITA	–2.69	–0.82	–0.24	–0.43	20.58	1.47	–15.83	–1.41
LX	0.41	1.08	0.37	1.47	0.23	0.11	0.34	0.06
NET	0.11	0.39	–0.05	–0.49	1.04	1.50	0.08	0.17
NOR	0.02	0.02	–0.07	–0.22	0.93	0.41	–4.22	–0.65
POR	–5.31***	–3.96	–0.65	–0.95	6.11***	2.75	–4.36**	–2.13
SWE	0.02	0.30	0.34***	6.37	0.90***	4.11	–0.89***	–5.16
SWI	–0.02	–0.09	0.22*	1.78	0.50	1.37	–1.28***	–2.88
TK	12.98	0.39	10.84	0.43	69.48	0.35	–46.29	–0.29
UK	0.18	0.35	0.40	1.21	2.47	0.61	–1.03	–0.64

*Notes: The symbols ***,**,* indicate levels of significance of 1%, 5%, and 10% respectively.*