

Brücker, Herbert; Siliverstovs, Boriss

**Working Paper**

## On the estimation and forecasting of international migration: how relevant is heterogeneity across countries?

IZA Discussion Papers, No. 1710

**Provided in Cooperation with:**

IZA – Institute of Labor Economics

*Suggested Citation:* Brücker, Herbert; Siliverstovs, Boriss (2005) : On the estimation and forecasting of international migration: how relevant is heterogeneity across countries?, IZA Discussion Papers, No. 1710, Institute for the Study of Labor (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/33516>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

IZA DP No. 1710

**On the Estimation and  
Forecasting of International Migration:  
How Relevant Is Heterogeneity Across Countries?**

Herbert Brücker  
Boriss Siliverstovs

August 2005

# On the Estimation and Forecasting of International Migration: How Relevant Is Heterogeneity Across Countries?

**Herbert Brücker**

*DIW Berlin and IZA Bonn*

**Boriss Siliverstovs**

*DIW Berlin*

Discussion Paper No. 1710  
August 2005

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
Email: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### On the Estimation and Forecasting of International Migration: How Relevant Is Heterogeneity Across Countries?\*

This paper performs a comparative analysis of estimation as well as of out-of-sample forecasting results of more than 20 estimators common in the panel data literature using the data on migration to Germany from 18 source countries in the period 1967-2001. Our results suggest that the choice of an estimation procedure has a substantial impact on the parameter estimates of the migration function. Out-of-sample forecasting results indicate the following: (i) the standard fixed effects estimators clearly outperforms the pooled OLS estimator, (ii) both the fixed effects estimators and the hierarchical Bayes estimator exhibit the superior forecast performance, (iii) the fixed effects estimators outperform GMM and other instrumental variables estimators, (iv) forecasting performance of heterogeneous estimators is mediocre in our data set.

JEL Classification: C23, C53, F22

Keywords: international migration, panel data, forecasting

Corresponding author:

Herbert Brücker  
DIW Berlin  
Königin-Luise-Straße 5  
D-14195 Berlin  
Germany  
Email: [hbruecker@diw.de](mailto:hbruecker@diw.de)

---

\* The authors wish to thank the Badi H. Baltagi, Paul Gregory Konstantin A. Kholodilin and the participants of the 'Aarhus Econometrics' seminar (Svinkløv, Denmark) for valuable comments and discussion of an earlier version of this paper. The authors are also indebted to Li Hongyi and A. Kamil Tahmiscioglu for providing the computer code for the shrinkage and hierarchical Bayes estimators, respectively. The usual disclaimer applies.

# 1 Introduction

International migration is a crucial policy issue. The scale of migration affects wages and employment opportunities in both the receiving and sending countries, and has important implications for public finance and the welfare state. In contrast to international goods and capital markets, international migration is effectively restricted in most parts of the world at present by legal and administrative barriers. However, political events such as the fall of the ‘iron curtain’ in Central and Eastern Europe and socio-economic developments such as demographic change and ageing have put the removal of migration barriers high on the agenda. The enlargement of the European Union (EU) is currently helping to lift restrictions to the free movement of workers and other persons from lower-income countries. Similarly, developed countries around the world such as Australia, Canada, and the US are considering lowering migration barriers to combat the negative effects of ageing and illegal migration.

The consequences of removing of migration barriers are, in many cases, irreversible once the policy decision has been made. Countries considering such a policy change thus urgently need reasonably accurate predictions of subsequent migration flows. Although international migration theory (see e.g. Burda, 1995; Stark, 1991; Stark and Taylor, 1991) has made considerable progress since the seminal contributions of Hicks (1932), Sjaastad (1962), and Harris and Todaro (1970), the empirical estimation of migration models still faces several methodological problems. The key problem to be addressed is heterogeneity across countries, which results mainly from differences in geography, history, language, and culture. Theoretical considerations suggest that these country-specific factors affect the costs and benefits of migration, and, hence, the scale of migration flows and stocks.

The heterogeneity across countries is addressed in different ways in the empirical migration literature. While some studies ignore country-specific effects completely, others consider time-invariant variables or capture individual-specific effects by estimating fixed effects models. As discussed in Alecke et al. (2001) and Brücker (2001), differences in estimation methodologies can lead to broadly divergent estimates of the migration models’ parameters. Consequently, the literature includes a vast array of estimates of future migration stocks and flows. For example, in anticipation of the upcoming EU enlargement, a number of studies have tried to assess the migration potential from the Central and Eastern European (CEE) accession countries to the incumbent EU members (for a survey see Straubhaar, 2002). Most studies expect that 2–4 percent of the CEE accession countries’ population will move to the fifteen incumbent EU member states (see e.g. Layard et al., 1992; Bauer and Zimmermann, 1999; Hille and Straubhaar, 2001; Boeri and Brücker et al., 2001). However some studies project substantially higher or lower figures. For instance, Sinn et al. (2001) and Flaig (2001) forecast the long-run migration potential from the CEE countries to Germany at 6–9 percent of the sending countries’ population, based on a pooled OLS model, while Brücker (2001) and Alvarez-Plata et al. (2003) estimate the long-run migration potential at 2–2.8 percent, based on a fixed effects model that otherwise shares the same features of the model employed by Sinn et al. (2001) and Flaig (2001) and uses a similar data set.

Clearly, differences among the parameter estimates of migration functions (and hence, of migration potentials) can also be attributed to the different data sets used, as well as to the different estimation methods employed. Typically, each study reports the parameters of a single specification of a migration function estimated by only one method. Hence it is often difficult to assess the extent to which the chosen estimation method influences the estimated parameters that underlie the forecasts of future migration potentials.

Against this background, the objective of this paper is twofold. First, we apply a variety of panel data estimators common in the econometric literature in order to determine how the parameter estimates differ depending on the method used. Second, we evaluate the out-of-sample forecasting performance of the different panel data estimators.

The approach applied here follows that of other studies, such as Baltagi and Griffin (1997), Baltagi et al. (2000, 2002, 2003, 2004), which compare the forecasting performance of panel data estimators in other contexts. The comparative analysis of various panel estimators allows us to address three questions in particular: First, do estimators that allow for country-specific effects outperform pooled estimators, which ignore these effects? Second, do estimators which address the problem of biased estimation of dynamic models in panel data sets with a limited time dimension (see Nickell, 1981; Kiviet, 1995), outperform traditional fixed effects and pooled estimators? Third, do estimators which consider heterogeneity not only in the intercept but also in the slope parameters (see Robertson and Symons, 1992; Pesaran and Smith, 1995; Hsiao et al., 1999; Maddala et al., 2001) outperform panel estimators which rely on the homogeneity assumption?

Our analysis is based on German migration data. We consider the migration to Germany from a panel of 18 European source countries during the period 1967–2001. The consideration of other European receiving countries is hampered by data limitations. Only a few EU members report migration flows and stocks for long time periods. Moreover, reporting systems and concepts of nationality differ among EU countries. Inconsistent definitions of migrants and measurement errors can therefore easily distort the data in panels of receiving countries. In view of these constraints, we limit our analysis to Germany. In contrast to most other EU countries, Germany reports annual data on migration flows and stocks by nationality since 1967, which makes it possible to use a panel with a relatively large time dimension. Note that Germany is the largest destination for immigrants within the present EU, with the share of approximately 40 percent, and, by far the largest destination of migration from the Central and Eastern European accession countries as well, with a 60 percent share. The data are described in the Appendix.

The remainder of this paper is structured as follows: in section 2 we briefly discuss the theoretical framework and specify the empirical model. In section 3, we describe different estimation procedures. In section 4 we present the estimation results. In section 5 we assess the forecasting performance. We summarize our major findings in the last section.

## 2 Model specification

The empirical analysis carried out here is based on a simple and fairly general migration model in the tradition of the so-called human capital approach, see Sjaastad (1962), Harris and Todaro (1970), Banerjee and Kanbur (1981), and Hatton (1995). Following this tradition, the macro-migration function is modelled as

$$mst_{ht} = f(w_{ft}/w_{ht}, w_{ht}, e_{ft}, e_{ht}, X_{ht}, Z_h), \quad (1)$$

where  $mst_{ht}$  is the share of migrants in percent of the home population,  $w$  is the wage measured at current exchange rates,  $e$  is the employment rate,  $X_{ht}$  is a vector of institutional variables which includes a dummy for the free movement within the EU and another dummy for bilateral guestworker agreements, and  $Z_h$  is a vector of country-specific time-invariant variables which includes the logarithm of the distance between the sending and the receiving country<sup>1</sup>, a dummy for geographical proximity, and a dummy for common language. The subscript  $f$  denotes the foreign country (Germany), the subscript  $h$  – the home country ( $h = 1, \dots, 18$ ), and the subscript  $t$  denotes the  $t$  –  $th$  year ( $t = 1967, \dots, 2001$ ).

In contrast to most other studies, we use migration *stocks* instead of migration *flows* as the dependent variable. This follows from the assumption that individuals are heterogeneous with respect to the human capital characteristics and preferences that affect the benefits and costs of migration. An aggregation of individual decisions thus leads to a declining propensity to migrate in the remaining population, the higher the share of the population already abroad.<sup>2</sup> Hence, a long-run equilibrium emerges between migration stocks and the explanatory variables rather than between migration flows and the explanatory variables. Thus, we understand net migration as a disequilibrium phenomenon, which eventually ceases when the equilibrium stock of migrants is achieved. This does not rule out the possibility that chain and network effects affect migration positively. But in the long run, these effects are dominated by declining preferences to migrate in the population.<sup>3</sup>

The choice of the explanatory variables is motivated by the following considerations: The decision to migrate is understood as an investment in human capital, whose returns are determined by the net present value of expected income streams in the future (Sjaastad, 1962). The costs of migration comprise not only the pecuniary costs of changing the place of residence, but also non-pecuniary costs including all social and psychological costs which result from moving to an unfamiliar environment. An individual will migrate if the expected benefits of moving exceed the expected costs.

Expectations on future income streams in the country of destination are conditioned on the opportunity to find a job on its labour market. Following Harris and Todaro (1970), the average employment rate – which is

---

<sup>1</sup>Following the literature on gravity models in trade, FDI and migration, distance is measured here as the distance between the capitals of the destination and the sending countries in km. For Germany we chose instead of the capital Frankfurt a.M. since (i) the German capital has moved from Bonn to Berlin during the sample period, and (ii) Frankfurt forms the geographical and economic centre of Germany.

<sup>2</sup>This is empirically proved for individual cohorts in the case of emigration from Norway by Baevre et al. (2001).

<sup>3</sup>For an analysis of the mechanics of migration stocks and flows in a model with heterogeneous agents see Brücker and Schröder (2005).

defined as the ratio of the employed labour force to the total labour force – serves as a proxy for the individual probability to find a job. Similar arguments apply to the expectations on future income in the home countries.

Following Faini and Venturini (1995), the wage variables enter the empirical model both as the ratio of the foreign to the home wage and as the home wage separately. The first variable captures the pecuniary incentives to migration resulting from the income differential. The inclusion of the second variable is motivated by liquidity constraints: At a given income differential between the receiving and the sending country, the number of migrants will increase with the home income since liquidity constraints are binding for a smaller share of the population the higher the income is in the sending country.

Finally, the choice of the institutional and the time-invariant variables is based on the following assumptions: First, we expect that the removal of institutional barriers reduces the pecuniary and social costs of migration. Hence, we expect that both the guestworker agreements between Germany and the sending countries, which introduced *de facto* free movement and supported the recruitment of labour actively by governmental institutions, and the free movement within the EU and the European Economic Area (EEA) increase the stock of migrants. Second, the consideration of time-invariant variables should capture both the geographical and the cultural distance between countries which affects the costs and benefits of migration. Geographical distance is measured here both in terms of km and with a dummy variable for geographical proximity. The former captures the fact that transport and communication costs increase with geographical distance, the latter that migrants tend to cluster in the first (large) destination country with a high per capita income. Moreover, we include a dummy variable for common language since language barriers are one of the main factors which determine the economic and social costs of migration.

We can consider of course a number of further variables which might affect the costs and benefits of migration as well. However, other time-invariant variables such as dummies for a common border, religion, and so on have turned out to be insignificant. Moreover, interactions between time-invariant and institutional variables on the one hand, and other explanatory variables such as the wage differential on the other hand, have been insignificant as well and are therefore not considered in the regressions reported here.

Given these considerations, the share of migrants who would like to reside in the foreign country under the given economic and other conditions is determined in the long-run equilibrium by:

$$mst_{ht}^* = \alpha^* + \beta_1^* \ln(w_{ft}/w_{ht}) + \beta_2^* \ln w_{ht} + \beta_3^* \ln e_{ft} + \beta_4^* \ln e_{ht} + X'_{ht} \gamma^* + Z'_h \lambda^*, \quad (2)$$

where the asteriks denotes the long-run (steady-state) values of the parameters. The semi-logarithmic functional form of the model follows from the assumption that the utility function is logarithmic, see Hatton (1995). Furthermore, we assume that migration is governed by a simple partial adjustment or habit-persistence model:

$$mst_{ht} - mst_{h,t-1} = \delta(mst_{ht}^* - mst_{h,t-1}) + u_{ht}, \quad (3)$$

where  $u_{ht}$  denotes the error term.



Substituting equation (2) into (3) yields the following dynamic migration model:

$$mst_{ht} = \alpha + (1 - \delta)mst_{h,t-1} + \beta_1 \ln(w_{ft}/w_{ht}) + \beta_2 \ln w_{ht} + \beta_3 \ln e_{ft} + \beta_4 \ln e_{ht} + X'_{ht}\gamma + Z'_h\lambda + u_{ht}, \quad (4)$$

where  $\alpha = \delta\alpha^*$ ,  $\gamma = \delta\gamma^*$ ,  $\lambda = \delta\lambda^*$ ,  $\beta_j = \delta\beta_j^*$  for  $j = 1, \dots, 5$ .

Finally, we follow convention and specify the error term as a one-way error-component model:<sup>4</sup>

$$u_{ht} = \mu_h + v_{ht},$$

where  $\mu_h$  denotes a country-specific effect, and  $v_{ht}$  is white noise.

From our theoretical considerations it follows that we expect the following signs for the long-run parameters of the model:  $\beta_1 > 0$ ,  $\beta_2 > 0$ ,  $\beta_3 > 0$ , and  $\beta_4 < 0$ . In addition, since migrants tend to be over-proportionally affected by employment risks in host countries, a larger coefficient for the employment rate in the host country is expected relative to the home country ( $\beta_3 + \beta_4 > 0$ ). For the coefficients on the dummy variables for free movement, guestworker agreements, geographical proximity, and common language we expect positive signs, for the coefficient on the distance variable a negative one.

The model in equation (4) forms the basis for our analysis. We assume that it is general enough to allow us to compare various estimators. In the empirical implementation, we allow one further lag of the endogenous variable in order to impose less restrictions on the adjustment process. Further lags of the dependent variable have turned out to be insignificant.

### 3 Discussion of the estimators

For our empirical analysis, we employ a number of the most widely used panel data estimators. The first set of estimators consists of the pooled OLS estimator without time-invariant variables (POLS), or with time-invariant variables (POLS with  $Z_h$ ). The former estimator completely ignores the individual-specific effects  $\mu_h$ , whereas the latter estimator ignores them only to the extent that they are not captured by  $Z_h$ . Since this estimator ignores country-specific effects (apart from those captured by  $Z_h$ ), the estimated parameters are certainly biased if the (omitted) country-specific effects are correlated with the lagged dependent and possibly some other explanatory variables.

The second group of estimators consists of those based on the ‘within’ transformation of the data that wipes out the time-invariant variables  $Z_h$ . All these estimators treat the country-specific effects  $\mu_h$  as fixed. We consider three fixed effects estimators, which impose different restrictions on the covariance matrix of the error term (Greene, 2002). The FE estimator assumes spherical disturbances. The FE(HET) estimator allows for the variance of residuals to differ across cross-sections (cross-sectional heteroscedasticity) and the FE(HET-COR) estimator allows both for cross-sectional heteroscedasticity and residual contemporaneous correlation across the cross-sections (cross-sectional correlation).

---

<sup>4</sup>The presence of the time-varying variables common to all cross-sections has precluded us from using the time dummies in our model due to the regressor multicollinearity.

The third group of estimators treats the country-specific effects  $\mu_h$  as random and therefore also employs variation between the cross-sections. We consider three versions of the random effects (RE) estimator: The Wallace and Hussain (1969), the Swamy and Arora (1972), and the iterated Feasible GLS (FGLS) estimator<sup>5</sup>, which apply different methods to weight optimally the 'within' and the 'between' variation in the data. Below we refer to these RE estimators as RE(WALHUS), RE(SWAR), and RE(MLE), respectively. Although all these RE estimators are asymptotically equivalent, in the relatively small sample used here, the estimated coefficients are likely to differ. Observe that random effects estimators allow us to include time-invariant variables  $Z_h$ . The corresponding versions are denoted by RE(WALHUS) with  $Z_h$ , RE(SWAR) with  $Z_h$ , and RE(MLE) with  $Z_h$ .

The fourth group of estimators addresses the problem of estimation bias which may result from the presence of the lagged dependent variable in dynamic panel models (Nickell, 1981; Kiviet, 1995). Although this bias disappears with the growing time dimension of the panel, it can still be relevant for the size of our panel with 33 observations over time (Judson and Owen, 1999). Therefore we consider the following estimators which address this issue: The pooled two-stage least squares (2SLS) estimator which uses the exogenous variables and their lags as instrumental variables without taking country specific effects into consideration, the FE-2SLS estimator which applies the within transformation and uses the exogenous variables and their lags as instruments, and the FD2SLS and FD2SLS-L estimators, which wipe out both the country-specific effects  $\mu_h$  and the time-invariant explanatory variables  $Z_h$  by transforming all variables in first-differences (Anderson and Hsiao, 1982). The FD2SLS estimator uses lags of the differenced variables and the FD2SLS-L lags of the variables in levels as instruments. Expanding on the Anderson and Hsiao (1982) estimators, Arellano and Bond (1991) show that there are many more instruments available within the GMM framework than used by conventional instrumental variable estimation. Consequently, the efficiency of GMM-based estimators is greatly enhanced<sup>6</sup>. However, as Baltagi et al. (2000) point out, the estimators of Anderson and Hsiao (1982) and Arellano and Bond (1991) may eliminate the estimation bias, but with a large loss of information. The application of the first-difference transformation destroys the economic structure formed between the levels of the variables across the time series dimension. Fortunately, the GMM estimator developed by Arellano and Bover (1995) addresses this issue by employing both the first differences as well as the levels of the variables by specifying the appropriate sets of instruments for both types of equations. We denote the Arellano and Bond (1991) and Arellano and Bover (1995) estimators as FDGMM and GMM, respectively, as the former is based on the first-difference transformation and the latter is based on the system of equations specified both for the first difference as well as for the levels of the variables. Note that the Arellano and Bover (1995) estimator also allows for inclusion of the time-invariant variables  $Z_h$  (GMM with  $Z_h$ ). Moreover, we report the estimation and the forecasting results for the one- and two-step GMM estimators (FDGMM1 and FDGMM2, GMM1 and GMM2, GMM1 with  $Z_h$  and GMM2 with  $Z_h$ ).

---

<sup>5</sup>The iterated FGLS estimator is equivalent to the Maximum Likelihood estimator.

<sup>6</sup>As noted in Phillips and Sul (2004), performance of these estimators can be unsatisfactory in situations where time series exhibit high persistence and weak instrumentation problems.

The fifth group represents the heterogeneous estimators. We consider the following estimators: the individual OLS (IOLS) estimator, the individual two-stage least squares (I2SLS) estimator, and the Mean Group (MG) estimator of Pesaran and Smith (1995). The first two estimators are based on the results of the individual country regressions, while the latter is based on the average parameter values of the individual regressions.

The sixth group of estimators contains the following shrinkage estimators: the empirical Bayes (Emp. Bayes), the iterative Bayes (Iter. Bayes), and the iterative empirical Bayes estimator (Iter. Emp. Bayes), see Maddala et al. (2001) for description. These shrinkage estimators constitute a compromise between the homogeneous estimators that restrict at least the slope to be uniform across different cross-sections and the heterogeneous estimators that allow the parameter sets to vary completely between every cross-section.

In addition to that, we apply the hierarchical Bayes (Hier. Bayes) estimator, see Hsiao et al. (1999). In contrast to all other estimators, the hierarchical Bayes estimator delivers the model parameter estimates using the Markov Chain Monte Carlo method (via Gibbs sampling). As reported in Hsiao et al. (1999) as well as Baltagi et al. (2004), this estimator has performed rather well both in Monte Carlo experiments and in out-of-sample forecasts.

## 4 Estimation Results

The estimated coefficient values for the short-run and long-run semi-elasticities are reported in Table 1, where for the sake of presentation, we report the sum of the coefficients of the first- and second lags of the migration stock variables with the associated standard error.

Before discussing the results for the individual estimators in detail, it is worthwhile to summarise some general observations: First, the sum of the autoregressive coefficients of the lagged dependent variable is below one for all estimators except the RE(SWAR) and some of the individual OLS and 2SLS estimators. Second, for a number of the estimators, the sum of the autoregressive coefficients is very close to one, implying a very high degree of persistence of the estimated model. This fact also results in rather high values of the long-run coefficient estimates, since the magnitude of the estimated long-run semi-elasticities becomes larger as the sum of the autoregressive coefficients approaches unity. Third, the estimated coefficients for the foreign-to-home wage ratio  $w_{ft}/w_{ht}$ , home wage  $w_{ht}$ , and German employment  $e_{ft}$  have the expected positive sign for almost all panel data estimators. However, the estimated coefficients for the home employment variable  $e_h$  have either positive or negative signs, depending on the estimation procedure, and are frequently insignificant. Fourth, the individual OLS and 2SLS regressions yield very heterogeneous coefficient estimates. Fifth, as expected, the shrinkage estimators substantially reduce the range of the coefficient estimates when compared to the individual regressions.

In the pooled OLS regressions (with and without the time-invariant variables  $Z_h$ ) the sum of the autoregressive coefficients is very close to unity, implying very high persistence in the dependent variable. Similar results are obtained also in other studies which employ the pooled OLS estimator not only in the migration

context (see e.g. Flaig, 2001; Sinn et al., 2001) but also for other data sets, e.g. cigarette demand in Baltagi et al. (2000), *inter alia*.

It is instructive to compare the results obtained for the POLS estimator with those obtained for the fixed effects estimators, which is the other popular type of panel data estimators that has been widely applied in the migration context, as noted above. As seen from Table 1, all three fixed effects estimators – FE, FE(HET), and FE(HET-COR) – yield much lower values of the sum of the autoregressive coefficients, which also results in much lower estimates of the long-run semi-elasticities.

The random effects estimators produce very heterogeneous results:<sup>7</sup> On the one hand, the results of the RE(WALHUS) estimates are very close to those of the pooled OLS. This can be explained by the fact that the calculation of the optimal weights for the RE(WALHUS) estimator is based on the OLS residuals, see Doornik et al. (2002). On the other hand, the results of the RE(MLE) regression are similar to those of the fixed effects estimates. This can be traced back to the fact that the importance of the ‘within’ variation increases with the growing time dimension in the GLS optimal weighting scheme, see Baltagi (2001). Finally, the sum of the autoregressive coefficients is above one in the RE(SWAR) estimate. The likely reason is that the RE(SWAR) estimator relies on the between regression in calculating the optimal weights for the ‘between’ and ‘within’ variation in the data. The relatively small number of the cross-sections ( $N = 18$ ) compared to the number of regressors (six without the time-invariant variables, and nine with the time-invariant variables) implies that there are only very few degrees of freedom left in the between regression.

We obtain also mixed results for the group of estimators which address the potential bias in dynamic panel estimation. On the one hand, the Anderson and Hsiao (1982) FD2SLS estimator and the Arellano and Bover (1995) GMM estimator with- and without the time-invariant variables  $Z_h$  uniformly imply a rather high persistence of the underlying time series, as the sum of the autoregressive coefficients is above 0.9. On the other hand, the Arellano and Bond (1991) FDGMM estimator yields estimates of the sum of the autoregressive coefficients below 0.8. It is also interesting to note that the estimates for the sum of the autoregressive coefficients in the pooled 2SLS regression with and without time invariant variables  $Z_h$  are comparable to those of the pooled OLS regression. At the same time, the estimation results of the FE-2SLS estimator are similar to those obtained by estimators that employ the ‘within’ transformation.

The individual OLS and 2SLS regressions yield very heterogeneous results. The coefficients for the sum of the lagged migration stock range from 0.182 to 1.038 in the individual OLS regressions, and from 0.182 to 1.087 in the individual 2SLS regressions, i.e. from models with a very low persistence over time to an explosive model. Moreover, the remaining slope coefficients in the individual regressions have either positive or negative signs, depending on the particular cross-section. The MG estimator yields for the sum of the autoregressive coefficients an estimate of 0.666, which is rather low in comparison to other estimators considered. It also obtains – contrary to our theoretical expectations – a negative sign on the foreign-to-home wage ratio.

---

<sup>7</sup>Since the inclusion of the time-invariant variables  $Z_h$  yields almost the same coefficient estimates as the models without  $Z_h$ , we limit our discussion to the random effects estimators without time-invariant variables here.

As expected, the shrinkage estimates of the sum of the autoregressive coefficient lie within a narrower range compared to those of the individual OLS and 2SLS regressions: 0.223 – 0.912 for the empirical Bayes estimator, 0.662 – 1.045 for the iterative Bayes estimator, and 0.783 – 1.004 for the iterative empirical Bayes estimator. Moreover, both iterative versions of the shrinkage estimators substantially narrow the range of the slope coefficients for the other explanatory variables in the model. Finally, the hierarchical Bayes estimator yields for the sum of the autoregressive coefficients an estimate of 0.779, which is somewhat lower than that of the estimators based on the ‘within’ transformation but it is close to the median values reported by the shrinkage estimators.

## 5 Forecasting Performance

For the evaluation of the out-of-sample forecasting performance of the different models, we have calculated the Root Mean Squared Error (RMSE), see Table 2. We compare the out-of-sample forecasting performance for two time periods: the fifth and the tenth year ahead. For this purpose, the estimates of the coefficients are based on the 1969–1996 and the 1969–1991 sample, respectively. The forecast accuracy is evaluated for the year 2001 in both forecast horizons. The results can be summarised as follows.

First, a comparison of the forecasting performance of the estimators that are widely applied in the migration literature, i.e. the pooled OLS and the fixed effects estimators, clearly demonstrates that the POLS (both with and without  $Z_h$ ) is inferior to that of the fixed effects estimator. As is evident from Table 2, the RMSE of the POLS estimators is around twice as high as that of the fixed effects estimators in both forecasting horizons.

Second, the fixed effects estimators as well as the hierarchical Bayes estimator show the lowest forecasting error for both time horizons. These results are comparable to those of a number of other studies including Baltagi and Griffin (1997) and Baltagi et al. (2000, 2002, 2003), where homogeneous estimators that allow for the individual-specific effects (e.g. fixed effects estimators) offer the best out-of-sample forecasting performance at long forecasting horizons. The superior forecasting performance of the hierarchical Bayes estimator has been recorded in Hsiao et al. (1999) as well as in the more recent study of Baltagi et al. (2004).

Third, the forecasting accuracy of the instrumental variables estimators, which address the problem of biased estimation in the dynamic panels, is well below that of the fixed effects estimator, which offer no correction for estimation bias.

Fourth, the iterative empirical Bayes estimator has shown by far the best performance of the three shrinkage estimators. Despite the fact that these shrinkage estimators outperform the heterogeneous OLS and 2SLS estimators, they perform comparatively worse than the homogeneous estimators based on the ‘within’ transformation. This can be attributed to their reliance upon the individual regression estimates, which show quite high degree of heterogeneity and parameter instability.

Finally, the mean group estimator of Pesaran and Smith (1995) shows the worst forecasting performance for both forecast horizons. This finding is similar to that reported in Baltagi et al. (2003) for a data set with

comparable dimensions to ours. Its poor performance can be explained by a number of causes: parameter-instability in the individual country regressions, the fact that averaging takes place only over 18 cross-sectional units, the low value of the estimated sum of the autoregressive coefficients, the unexpected (negative) sign of the estimated coefficient for the wage-ratio of the receiving and the sending country, and, finally, the fact that country-specific effects have not been considered when producing forecasts.

## 6 Conclusions

Economic intuition suggests that country-specific factors such as language, geography, history, and culture affect the benefits and costs of migration, and, hence, the scale of migration flows and stocks. This heterogeneity across countries can affect both the intercepts and the slope parameters of macro migration models. Large parts of the applied research devoted to predicting aggregate migration potentials ignore however country-specific effects. In this study, we have compared a wide variety of homogeneous and heterogeneous panel estimators for the purpose of assessing their out-of-sample forecasting performance.

We find that the fixed effects estimators and the hierarchical Bayes estimator offer the highest forecast accuracy at both forecasting horizons in comparison to all other estimators considered in this study, including the pooled OLS estimator, random effects estimators, heterogeneous estimators, instrumental variable estimators, and shrinkage estimators. The forecasting error of the pooled OLS estimator, which is due to its simplicity widely applied in the empirical migration literature, is around twice as high as that of the fixed effects estimators in both forecasting periods.

The finding that simple fixed effects estimators are a robust forecasting device for international migration in panel data with a relatively short time dimension has important quantitative consequences for migration forecasts: The widely applied alternative, the pooled OLS estimator, yields much larger estimates for the coefficients on the lagged dependent variable, and, consequently, larger long-run elasticities between migration stocks and the explanatory variables. Consequently, the estimates of migration potentials obtained by pooled OLS estimators are much larger than those obtained by fixed effects estimators. As an example, in the context of the Eastern Enlargement of the EU, pooled OLS estimators predict for Germany a long-run migration potential from the Central and Eastern Europe which is 2-3 times larger than that predicted by fixed effects models.

The other important finding of this study is that heterogeneous models perform poorly relative to homogeneous estimators. Moreover, shrinkage estimators, which reduce the range of the parameter estimates relative to the individual regressions, and are therefore recommended in cases where the variance of the individual results is high, show in our data set only a mediocre forecasting performance. These findings are at first glance surprising, since it is reasonable to assume that heterogeneity across countries affects not only the intercept, but also the slope parameters of migration models. A possible explanation for this seeming paradox is that the results of the individual regressions are highly unstable, such that the higher efficiency from pooling may more than offset the biases which result from inter-country heterogeneity. Note that our findings confirm the results

of previous research obtained in different contexts such as gasoline demand, cigarette demand, and electricity and natural gas consumption (see Baltagi and Griffin, 1997; Baltagi et al., 2000, 2002, 2003, respectively).

Last but not least, our study further reinforces the findings of yet rather small but growing literature that emphasizes superior forecasting performance of the hierarchical Bayes estimator suggested in Hsiao et al. (1999).

Our findings apply of course only to the data set used here. Nevertheless, Germany is the largest destination of migration in Europe and one of the largest in the world. Future research has to show whether the results presented here are robust if other data sets are employed.

## References

- Alecke, B., P. Huber, and G. Untiedt (2001). What a difference a constant makes. How predictable are international migration flows? In OECD (Ed.), *Migration Policies and EU-Enlargement, the Case of Central and Eastern Europe*, pp. 63–78. OECD, Paris.
- Alvarez-Plata, P., H. Brücker, and B. Siliverstovs (2003). Potential migration from Central and Eastern Europe into the EU-15 - an update. Report for the European Commission, DG Employment and Social Affairs, Berlin: German Institute for Economic Research (DIW Berlin).
- Anderson, T. W. and C. Hsiao (1982). Estimation of dynamic panel models with error components. *Journal of the American Statistical Association* 76, 74–82.
- Arellano, M. and S. R. Bond (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* 58, 277–97.
- Arellano, M. and O. Bover (1995). Another look at the instrumental-variable estimation of error-components models. *Journal of Econometrics* 68, 29–52.
- Baevre, K., C. Riis, and T. Thonstad (2001). Norwegian cohort emigration. *Journal of Population Economics* 14(3), 473–489.
- Baltagi, B. (2001). *Econometric Analysis of Panel Data* (2nd ed.). Chichester: John Wiley and Sons.
- Baltagi, B. H., G. Bresson, J. M. Griffin, and A. Pirotte (2003). Homogeneous, heterogenous or shrinkage estimators? Some empirical evidence from French regional gasoline consumption. *Empirical Economics* 28, 795–811.
- Baltagi, B. H., G. Bresson, and A. Pirotte (2002). Comparison of forecast for homogeneous, heterogeneous and shrinkage estimators. Some empirical evidence for US electricity and natural-gas consumption. *Economics Letters* 76, 375–382.
- Baltagi, B. H., G. Bresson, and A. Pirotte (2004). Tobin q: Forecast performance for hierarchical Bayes, shrinkage, heterogeneous and homogeneous panel data estimators. *Empirical Economics* 29, 107–113.
- Baltagi, B. H. and J. M. Griffin (1997). Pooled estimators vs. their heterogeneous counterparts in the context of dynamic demand for gasoline. *Journal of Econometrics* 77, 303–327.
- Baltagi, B. H., J. M. Griffin, and W. Xiong (2000). To pool or not to pool: Homogeneous versus heterogeneous estimators applied to cigarette demand. *The Review of Economics and Statistics* 82(1), 117–126.
- Banerjee, B. and S. M. R. Kanbur (1981). On the specification and estimation of macro rural-urban migration functions with an application to Indian data. *Oxford Bulletin of Economics and Statistics* 43(1), 7–29.



- Bauer, T. and K. F. Zimmermann (1999). Assessment of possible migration pressure and its labour market impact following EU enlargement to Central and Eastern Europe. Bonn: Institute for the Study of Labor (IZA).
- Boeri, T. and H. Brücker et al. (2001). The impact of eastern enlargement on employment and labour markets in the EU member states. Report for European Commission, DG Employment and Social Affairs, Brussels.
- Brücker, H. (2001). Die Folgen der Freizügigkeit für die Ost-West-Migration. Schlussfolgerungen aus einer Zeitreihenanalyse der Migration nach Deutschland, 1967 bis 1998. *Konjunkturpolitik* 52(supplement), 17–54.
- Brücker, H. and P. J. H. Schröder (2005). International migration with heterogeneous agents: Theory and evidence. Mimeo, DIW Berlin.
- Burda, M. C. (1995). Migration and the option value of waiting. *Economic and Social Review* 27(1), 1–19.
- Doornik, J. A., M. Arellano, and S. R. Bond (2002). Panel data estimation using DPD for Ox. Manual, available online: <http://www.nuff.ox.ac.uk//users//doornik>.
- Faini, R. and A. Venturini (1995). Migration and growth: The experience of southern Europe. CEPR discussion paper No. 964.
- Flaig, G. (2001). Zur Abschätzung der Migrationspotentiale der osteuropäischen EU-beitrittsländer. *Konjunkturpolitik Supplement* 52, 55–76.
- Greene, W. H. (2002). *Econometric Analysis* (5th ed.). Prentice Hall.
- Harris, J. and M. Todaro (1970). Migration, unemployment, and development: A two-sector analysis. *American Economic Review* 60, 126–142.
- Hatton, T. J. (1995). A model of UK migration 1870-1913. *Review of Economics and Statistics* 77(3), 407–415.
- Hicks, J. (1932). *The Theory of Wages*. London: McMillan.
- Hille, H. and T. Straubhaar (2001). The impact of EU -enlargement on migration movements and economic integration: Results of recent studies. In OECD (Ed.), *Migration Policies and EU-Enlargement, the Case of Central and Eastern Europe*, pp. 79–100. OECD, Paris.
- Hsiao, C., M. H. Pesaran, and A. K. Tahmiscioglu (1999). Bayes estimation of short-run coefficients in dynamic panel data models. In C. Hsiao, K. Lahiri, L.-F. Lee, and M. H. Pesaran (Eds.), *Analysis of Panels and Limited Dependent Variable Models*, pp. 268–296. Cambridge: Cambridge University Press.
- Judson, R. A. and A. K. Owen (1999). Estimating dynamic panel data models: A guide for macroeconomists. *Economics Letters* 65, 9–15.

- Kiviet, J. F. (1995). On bias, inconsistency and efficiency of some estimators in dynamic panel data models. *Journal of Econometrics* 68, 53–78.
- Layard, R., O. J. Blanchard, R. Dornbusch, and P. Krugman (1992). *East-West Migration: The Alternatives*. Cambridge, MA: MIT-Press.
- Maddala, G. S., H. Li, and V. K. Srivastava (2001). A comparative study of different shrinkage estimators for panel data models. *Annals of Economics and Finance* 2, 1–30.
- Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica* 49, 1417–1426.
- Pesaran, H. M. and R. Smith (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics* 68, 79–113.
- Phillips, P. C. B. and D. Sul (2004). Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence. Cowles Foundation Discussion Paper No. 1438.
- Robertson, D. and J. Symons (1992). Some strange properties of panel data estimators. *Journal of Applied Econometrics* 7, 175–189.
- Sinn, H.-W., G. Flaig, M. Werding, S. Munz, N. Düll, and H. Hoffmann (2001). EU-Erweiterung und Arbeitskräftemigration. Wege zu einer schrittweisen Annäherung der Arbeitsmärkte. München: Ifo-Institut für Wirtschaftsforschung.
- Sjaastad, L. (1962). The costs and returns of human migration. *Journal of Political Economy* 70, 80–93.
- Stark, O. (1991). *The Migration of Labour*. Cambridge, MA: Basil Blackwell.
- Stark, O. and J. E. Taylor (1991). Migration incentives, migration types: The role of relative deprivation. *Economic Journal* 101(408), 1163–1178.
- Straubhaar, T. (2002). Ost-West Migrationspotenzial: Wie gross ist es? *Jahrbücher für Nationalökonomie und Statistik* 222(1), 21–42.
- Swamy, P. A. V. B. and S. S. Arora (1972). The exact finite sample properties of the estimators of coefficients in the error component regression models. *Econometrica* 40, 261–275.
- Wallace, T. D. and A. Hussain (1969). The use of error components model in combining cross-section and time-series data. *Econometrica* 37, 55–72.

## 7 Appendix: Data Description

The sample used for the econometric analysis in this paper contains 18 source countries (Austria, Belgium, Denmark, Finland, France, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, United Kingdom). This sample covers the European source countries of migration to Germany almost completely, with the exception of the countries of the former COMECON and the (former) Yugoslavia. The COMECON countries have been excluded since the ‘iron curtain’ effectively prevented migration for the main period of analysis, and the former Yugoslavia has been excluded as well since the civil wars have heavily affected migration from there.

The dependent variable is the share of foreign citizens residing in Germany as a percentage of the home population. Foreign nationals are defined by their citizenship. Note that citizenship is granted on basis of the concept of ethnicity in Germany, such that the large majority of second- and third-generation migrants still possess foreign citizenship. Data on the foreign-born population are not available in the German statistics. The data on foreign residents stem from the Federal Statistical Office (Statistisches Bundesamt, Fachserie 1). Foreign residents have been reported in Germany since 1967 on an annual basis by the local municipalities, and have been counted by the central register of foreign nationals (Ausländerzentralregister) in Cologne since 1972. In general, the foreigner statistics in Germany tend to overreport the number of legal migrants slightly, since return migration is not completely recorded in the official figures.

In the sample period, we observe two statistical breaks: First, the transition of paper-based counting of foreign nationals by the local municipalities to computer-based counting by the central register of foreigners in 1972 produced a minor statistical break in case of some source countries (Statistisches Bundesamt, 1999, p. 5). The second break emerged after a revision of the foreigner statistics in the course of the population census of 1987, which reduced foreigner figures significantly for a period of three years (Statistisches Bundesamt, 1989, p. 594). After three years, the statistics were based again on the non-revised figures of the central register of foreigners, however. In order to control for the first break, we included a dummy variable in the regressions, but this turned out to be insignificant. We thus decided to ignore this break. With respect to the second break, we recalculated the number of foreign residents on the basis of net migration figures for the three years affected by the revisions of the Federal Statistical Office.

The migration stock variable is normalized by the population of the home countries. Population figures are taken from the World Bank (2002). The dependent variable in the econometric analysis is the change in the migration stock as a percentage of the home population. By definition, this deviates from the net migration rate by the rate of natural population growth of the migrant population relative to that of the home population and the rate of naturalisations, i.e.:

$$\Delta mst_t = m_t + \frac{n_{ft} - n_{ht} - \delta}{1 + n_{ht}} mst_{t-1}, \quad (5)$$

where  $mst$  is defined as the ratio of the stock of residents to the home population,  $m$  as the ratio of net migration

to the home population,  $n_f$  as the rate of natural population growth in the migrant population,  $n_h$  as the rate of natural population growth in the home population,  $\delta$  the rate of naturalisations in the migrant population, and  $t$  the time index. Thus, the change in the migrant stock equals the net migration rate if the rate of natural population growth in the migrant population equals the sum of the rate of natural population growth in the home country and the rate of naturalisations. In our sample, the difference between the net migration rate and the change in migration stocks is moderate.

The explanatory variables in our model are per capita income and employment rates in Germany and the source countries. Consistent wage variables are not available for our country sample. Following the literature we thus use GDP per capita measured at purchasing power parity as an approximation for the income level. The per capita GDP at purchasing power parity (PPP-GDP) series is taken from Maddison (1995) for the period 1967-1994 and has been extrapolated with the real growth rate of the PPP-GDP per capita. The latter has been taken again from the OECD Main Economic Indicators and Historical Statistics and is complemented by national sources for countries not covered by the OECD series.

The employment rate in the econometric analysis is calculated as one minus the unemployment rate. The ILO-definition for the unemployment rates have been used; time series for the unemployment rates stem from the OECD and are complemented by national statistical sources.

Table 1: Regression Results: Short- and long-run semi-elasticities

|                       |                 | Short-run                          |                      |           |              |           | Long-run             |           |              |           |  |
|-----------------------|-----------------|------------------------------------|----------------------|-----------|--------------|-----------|----------------------|-----------|--------------|-----------|--|
|                       |                 | $mst_{h,t-1}^1$<br>& $mst_{h,t-2}$ | $\ln(w_{ft}/w_{ht})$ | $\ln w_h$ | $\ln e_{ft}$ | $\ln e_h$ | $\ln(w_{ft}/w_{ht})$ | $\ln w_h$ | $\ln e_{ft}$ | $\ln e_h$ |  |
| POLS                  | Coeff.          | 0.991                              | 0.106                | 0.090     | 0.608        | -0.002    | 11.635               | 9.837     | 66.636       | -0.245    |  |
|                       | <i>t</i> -value | 179.57                             | 4.05                 | 3.74      | 2.69         | -0.04     | 1.8                  | 1.56      | 1.28         | -0.04     |  |
| POLS with $Z_h$       | Coeff.          | 0.977                              | 0.134                | 0.096     | 0.585        | 0.019     | 5.768                | 4.153     | 25.244       | 0.83      |  |
|                       | <i>t</i> -value | 135.88                             | 4.67                 | 4.05      | 2.83         | 0.37      | 2.98                 | 3         | 2.17         | 0.36      |  |
| FE                    | Coeff.          | 0.871                              | 0.105                | 0.117     | 0.687        | -0.096    | 0.818                | 0.912     | 5.338        | -0.749    |  |
|                       | <i>t</i> -value | 32.737                             | 2.85                 | 4.72      | 3.97         | -1.36     | 2.23                 | 3.2       | 2.71         | -1.31     |  |
| FE(HET)               | Coeff.          | 0.896                              | 0.044                | 0.041     | 0.217        | -0.062    | 0.426                | 0.398     | 2.085        | -0.592    |  |
|                       | <i>t</i> -value | 83.25                              | 18.54                | 22.33     | 14.14        | -10.06    | 9.21                 | 13.2      | 6.93         | -6.62     |  |
| FE(HET,COR)           | Coeff.          | 0.874                              | 0.082                | 0.091     | 0.478        | -0.084    | 0.649                | 0.722     | 3.784        | -0.661    |  |
|                       | <i>t</i> -value | 140                                | 7.92                 | 8.96      | 7.62         | -12.01    | 7.93                 | 8.98      | 7.16         | -10.71    |  |
| RE(WALHUS)            | Coeff.          | 0.989                              | 0.11                 | 0.091     | 0.622        | -0.022    | 9.805                | 8.107     | 55.54        | -1.968    |  |
|                       | <i>t</i> -value | 255.82                             | 3.31                 | 2.93      | 3.11         | -0.29     | 2.52                 | 2.13      | 2.02         | -0.28     |  |
| RE(WALHUS) with $Z_h$ | Coeff.          | 0.97                               | 0.143                | 0.099     | 0.608        | -0.012    | 4.73                 | 3.281     | 20.1         | -0.397    |  |
|                       | <i>t</i> -value | 171.32                             | 4.08                 | 3.20      | 3.08         | -0.15     | 3.75                 | 2.86      | 2.57         | -0.15     |  |

*continued overleaf*

Table 1: Short- and long-run semi-elasticities (continued)

|                     |            | Short-run                          |                      |           |              |           | Long-run             |           |              |           |  |
|---------------------|------------|------------------------------------|----------------------|-----------|--------------|-----------|----------------------|-----------|--------------|-----------|--|
|                     |            | $mst_{h,t-1}^1$<br>& $mst_{h,t-2}$ | $\ln(w_{ft}/w_{ht})$ | $\ln w_h$ | $\ln e_{ft}$ | $\ln e_h$ | $\ln(w_{ft}/w_{ht})$ | $\ln w_h$ | $\ln e_{ft}$ | $\ln e_h$ |  |
| RE(SWAR)            | Coeff.     | 1.001                              | 0.071                | 0.063     | 0.567        | 0.056     | na                   | na        | na           | na        |  |
|                     | $t$ -value | 1742.6                             | 2.10                 | 1.87      | 2.68         | 3.63      | na                   | na        | na           | na        |  |
| RE(SWAR) with $Z_h$ | Coeff.     | 1.001                              | 0.060                | 0.051     | 0.496        | 0.047     | na                   | na        | na           | na        |  |
|                     | $t$ -value | 2028.6                             | 1.73                 | 1.48      | 2.30         | 4.37      | na                   | na        | na           | na        |  |
| RE(MLE)             | Coeff.     | 0.896                              | 0.150                | 0.116     | 0.678        | -0.095    | 1.451                | 1.12      | 6.544        | -0.919    |  |
|                     | $t$ -value | 85.68                              | 3.39                 | 3.95      | 3.53         | -0.92     | 3.33                 | 3.82      | 3.21         | -0.90     |  |
| RE(MLE) with $Z_h$  | Coeff.     | 0.896                              | 0.147                | 0.116     | 0.672        | -0.087    | 1.405                | 1.109     | 6.429        | -0.833    |  |
|                     | $t$ -value | 85.27                              | 3.33                 | 3.93      | 3.49         | -0.86     | 3.26                 | 3.81      | 3.19         | -0.83     |  |
| 2SLS                | Coeff.     | 0.947                              | 0.231                | 0.134     | 0.558        | 0.269     | 4.353                | 2.526     | 10.538       | 5.075     |  |
|                     | $t$ -value | 28.75                              | 2.92                 | 3.26      | 1.61         | 1.26      | 2.64                 | 1.65      | 0.89         | 3.08      |  |
| 2SLS with $Z_h$     | Coeff.     | 0.972                              | 0.171                | 0.120     | 0.770        | 0.039     | 6.181                | 4.343     | 27.766       | 1.423     |  |
|                     | $t$ -value | 40.93                              | 3.29                 | 3.37      | 2.88         | 0.35      | 1.44                 | 1.18      | 0.94         | 0.44      |  |
| FE-2SLS             | Coeff.     | 0.872                              | 0.110                | 0.122     | 0.757        | -0.122    | 0.857                | 0.951     | 5.899        | -0.951    |  |
|                     | $t$ -value | 18.60                              | 2.26                 | 3.83      | 2.78         | -0.75     | 1.68                 | 2.57      | 1.60         | -0.62     |  |
| FD2SLS              | Coeff.     | 0.933                              | 0.201                | 0.324     | 1.074        | -0.142    | 2.992                | 4.811     | 15.956       | -2.111    |  |

*continued overleaf*

Table 1: Short- and long-run semi-elasticities (continued)

|                 |                 | Short-run                          |                      |           |              | Long-run  |                      |           |              |           |
|-----------------|-----------------|------------------------------------|----------------------|-----------|--------------|-----------|----------------------|-----------|--------------|-----------|
|                 |                 | $mst_{h,t-1}^1$<br>& $mst_{h,t-2}$ | $\ln(w_{ft}/w_{ht})$ | $\ln w_h$ | $\ln e_{ft}$ | $\ln e_h$ | $\ln(w_{ft}/w_{ht})$ | $\ln w_h$ | $\ln e_{ft}$ | $\ln e_h$ |
|                 | <i>t</i> -value | 2.36                               | 0.82                 | 1.83      | 3.27         | -0.55     | 0.14                 | 0.16      | 0.17         | -0.15     |
| FD2SLS-L        | Coeff.          | 0.975                              | 0.13                 | 0.212     | 0.984        | -0.062    | 5.121                | 8.363     | 38.756       | -2.444    |
|                 | <i>t</i> -value | 12.48                              | 0.75                 | 4.14      | 2.77         | -0.24     | 0.26                 | 0.32      | 0.32         | -0.20     |
| FDGMM1          | Coeff.          | 0.765                              | 0.049                | 0.133     | 1.188        | -0.225    | 0.208                | 0.567     | 5.049        | -0.958    |
|                 | <i>t</i> -value | 29.54                              | 0.23                 | 2.20      | 3.20         | -0.72     | 0.23                 | 1.93      | 2.95         | -0.71     |
| FDGMM2          | Coeff.          | 0.793                              | 0.092                | 0.124     | 1.065        | 0.026     | 0.445                | 0.598     | 5.143        | 0.127     |
|                 | <i>t</i> -value | 21.18                              | 0.63                 | 2.35      | 3.16         | 0.13      | 0.60                 | 2.10      | 2.72         | 0.13      |
| GMM1            | Coeff.          | 0.977                              | 0.167                | 0.127     | 0.811        | 0.036     | 7.373                | 5.573     | 35.701       | 1.604     |
|                 | <i>t</i> -value | 110.04                             | 4.22                 | 4.09      | 2.58         | 0.48      | 3.49                 | 2.52      | 1.63         | 0.47      |
| GMM2            | Coeff.          | 0.981                              | 0.154                | 0.122     | 0.665        | 0.045     | 7.899                | 6.229     | 34.117       | 2.309     |
|                 | <i>t</i> -value | 111.7                              | 3.82                 | 3.85      | 2.21         | 0.41      | 3.09                 | 2.52      | 1.51         | 0.39      |
| GMM1 with $Z_h$ | Coeff.          | 0.950                              | 0.212                | 0.138     | 0.772        | 0.073     | 4.210                | 2.748     | 15.317       | 1.445     |
|                 | <i>t</i> -value | 65.38                              | 5.40                 | 4.42      | 2.62         | 0.82      | 3.51                 | 3.48      | 2.00         | 0.80      |
| GMM2 with $Z_h$ | Coeff.          | 0.960                              | 0.162                | 0.115     | 0.840        | -0.097    | 4.029                | 2.862     | 20.936       | -2.420    |
|                 | <i>t</i> -value | 42.52                              | 2.51                 | 2.37      | 2.68         | -0.59     | 1.86                 | 1.88      | 1.26         | -0.48     |

*continued overleaf*

Table 1: Short- and long-run semi-elasticities (continued)

|       |            | Short-run                          |                      |           |              | Long-run  |                      |           |              |           |
|-------|------------|------------------------------------|----------------------|-----------|--------------|-----------|----------------------|-----------|--------------|-----------|
|       |            | $mst_{h,t-1}^1$<br>& $mst_{h,t-2}$ | $\ln(w_{ft}/w_{ht})$ | $\ln w_h$ | $\ln e_{ft}$ | $\ln e_h$ | $\ln(w_{ft}/w_{ht})$ | $\ln w_h$ | $\ln e_{ft}$ | $\ln e_h$ |
| iOLS  |            |                                    |                      |           |              |           |                      |           |              |           |
| Max   | Coeff.     | 1.038                              | 0.811                | 0.442     | 5.835        | 1.979     | 4.555                | 3.379     | 35.922       | 11.075    |
|       | $t$ -value | 21.17                              | 0.71202              | 3.2839    | 2.1504       | 0.71184   | 0.68                 | 1.00      | 0.85         | 0.55      |
| Med   | Coeff.     | 0.711                              | 0.044                | 0.042     | 0.427        | -0.286    | 0.120                | 0.175     | 1.234        | -0.854    |
|       | $t$ -value | 10.23                              | 0.79                 | 2.13      | 1.79         | -2.69     | 0.82                 | 0.49      | 1.28         | -2.26     |
| Min   | Coeff.     | 0.182                              | -2.799               | -0.173    | -0.581       | -4.044    | -5.875               | -0.912    | -0.797       | -22.708   |
|       | $t$ -value | 0.87                               | -1.81                | -0.74     | -0.69        | -1.57     | -3.63                | -0.34     | -0.72        | -1.38     |
| MG    | Coeff.     | 0.666                              | -0.125               | 0.089     | 1.23         | -0.451    | -0.375               | 0.267     | 3.682        | -1.35     |
|       | $t$ -value | 11.09                              | -1.03                | 2.37      | 5.18         | -1.56     | -1.01                | 2.19      | 3.79         | -1.50     |
| i2SLS |            |                                    |                      |           |              |           |                      |           |              |           |
| Max   | Coeff.     | 1.087                              | 1.036                | 0.401     | 5.497        | 1.933     | 27.678               | 20.903    | 102.250      | 3.055     |
|       | $t$ -value | 4.37                               | 0.82                 | 2.44      | 1.87         | 0.69      | 0.15                 | 0.15      | 0.15         | 0.65      |
| Med   | Coeff.     | 0.756                              | 0.064                | 0.046     | 0.426        | -0.183    | 0.368                | 0.209     | 1.048        | -0.542    |
|       | $t$ -value | 5.13                               | 1.39                 | 1.95      | 1.44         | -2.02     | 1.33                 | 0.61      | 1.35         | -1.90     |
| Min   | Coeff.     | 0.065                              | -3.781               | -0.204    | -0.432       | -4.303    | -6.028               | -1.656    | -9.649       | -34.870   |

*continued overleaf*



Table 1: Short- and long-run semi-elasticities (continued)

|             |                 | Short-run                        |                      |           |              | Long-run  |                      |           |              |           |
|-------------|-----------------|----------------------------------|----------------------|-----------|--------------|-----------|----------------------|-----------|--------------|-----------|
|             |                 | $mst_{h,t-1}^1$                  | $\ln(w_{ft}/w_{ht})$ | $\ln w_h$ | $\ln e_{ft}$ | $\ln e_h$ | $\ln(w_{ft}/w_{ht})$ | $\ln w_h$ | $\ln e_{ft}$ | $\ln e_h$ |
|             |                 | <i>&amp; mst<sub>h,t-2</sub></i> |                      |           |              |           |                      |           |              |           |
|             | <i>t</i> -value | 0.15                             | -1.46                | -0.41     | -0.41        | -1.45     | -3.47                | -0.37     | -0.35        | -0.89     |
| Emp. Bayes  |                 |                                  |                      |           |              |           |                      |           |              |           |
| Max         | Coeff.          | 0.912                            | 0.468                | 0.468     | 6.392        | 1.406     | 3.627                | 2.406     | 33.861       | 2.319     |
|             | <i>t</i> -value | 7.38                             | 2.45                 | 6.85      | 5.92         | 3.25      | 0.79                 | 0.74      | 0.95         | 3.01      |
| Med         | Coeff.          | 0.753                            | 0.055                | 0.043     | 0.623        | -0.350    | 0.125                | 0.169     | 2.007        | -1.310    |
|             | <i>t</i> -value | 9.43                             | 1.16                 | 3.51      | 2.39         | -4.07     | 0.58                 | 3.11      | 1.27         | -0.89     |
| Min         | Coeff.          | 0.223                            | -1.987               | -0.321    | -0.029       | -4.689    | -2.700               | -1.700    | -0.123       | -24.838   |
|             | <i>t</i> -value | 1.24                             | -5.27                | -2.77     | -0.12        | -5.81     | -2.38                | -1.01     | -0.13        | -0.99     |
| Iter. Bayes |                 |                                  |                      |           |              |           |                      |           |              |           |
| Max         | Coeff.          | 1.045                            | 0.143                | 0.123     | 2.019        | -0.071    | 1.794                | 1.093     | 9.789        | -0.224    |
|             | <i>t</i> -value | 32.87                            | 2.68                 | 5.50      | 8.52         | -2.09     | 1.36                 | 1.57      | 4.65         | -1.62     |
| Med         | Coeff.          | 0.812                            | 0.033                | 0.060     | 0.389        | -0.133    | 0.091                | 0.308     | 1.282        | -0.725    |
|             | <i>t</i> -value | 22.23                            | 1.56                 | 6.87      | 6.61         | -5.45     | 0.87                 | 5.01      | 1.38         | -3.86     |
| Min         | Coeff.          | 0.662                            | -0.172               | 0.009     | 0.044        | -0.207    | -0.835               | 0.049     | 0.141        | -3.263    |
|             | <i>t</i> -value | 8.14                             | -2.81                | 1.22      | 0.18         | -6.10     | -2.81                | 1.04      | 0.17         | -1.27     |

*continued overleaf*

Table 1: Short- and long-run semi-elasticities (continued)

| Iter. Emp Bayes | $mst_{h,t-1}^1$<br>& $mst_{h,t-2}$ | Short-run            |           |              |           | Long-run             |           |              |           |        |
|-----------------|------------------------------------|----------------------|-----------|--------------|-----------|----------------------|-----------|--------------|-----------|--------|
|                 |                                    | $\ln(w_{ft}/w_{ht})$ | $\ln w_h$ | $\ln e_{ft}$ | $\ln e_h$ | $\ln(w_{ft}/w_{ht})$ | $\ln w_h$ | $\ln e_{ft}$ | $\ln e_h$ |        |
| Max             | Coeff.                             | 1.004                | 0.086     | 0.085        | 1.183     | -0.005               | 1.102     | 0.747        | 6.431     | -0.028 |
|                 | $t$ -value                         | 28.27                | 2.60      | 5.05         | 7.07      | -0.13                | 1.57      | 2.25         | 4.23      | -0.13  |
| Med             | Coeff.                             | 0.830                | 0.044     | 0.052        | 0.312     | -0.115               | 0.238     | 0.345        | 2.248     | -0.588 |
|                 | $t$ -value                         | 29.92                | 1.99      | 13.72        | 6.56      | -5.88                | 0.92      | 5.78         | 1.71      | -2.89  |
| Min             | Coeff.                             | 0.783                | 0.001     | 0.011        | 0.186     | -0.148               | 0.003     | 0.065        | 1.222     | -2.056 |
|                 | $t$ -value                         | 15.77                | 0.02      | 1.75         | 2.39      | -4.02                | 0.015     | 1.46         | 1.84      | -1.84  |
| Hier. Bayes     | Coeff.                             | 0.779                | 0.018     | 0.071        | 1.028     | -0.334               | 0.081     | 0.323        | 4.653     | -1.512 |
|                 | $t$ -value                         | 15.92                | 0.12      | 1.60         | 2.30      | -1.12                | 0.12      | 1.54         | 1.92      | -1.06  |

The  $t$ -values of the GMM 2-step estimators are based on standard errors obtained after small sample correction suggested in Windmeijer (2000). For description of the estimators see Section 3.

<sup>1</sup> This column contains sum of autoregressive coefficients for the  $mst_{h,t-1}$  and  $mst_{h,t-2}$  variables.

Table 2: Comparison of forecasting performance

| Rank                  | 5 years ahead |                       | 10 years ahead |
|-----------------------|---------------|-----------------------|----------------|
| Estimator             | RMSE          | Estimator             | RMSE           |
| FE(HET-COR)           | 0.095         | Hier. Bayes           | 0.112          |
| FE                    | 0.098         | FE(HET)               | 0.112          |
| FE-2SLS               | 0.098         | FE(HET-COR)           | 0.148          |
| FE(HET)               | 0.106         | FE-2SLS               | 0.149          |
| Hier. Bayes           | 0.107         | FE                    | 0.155          |
| Iter. Emp. Bayes      | 0.111         | FD2SLS                | 0.180          |
| RE(MLE)               | 0.145         | FDGMM1                | 0.185          |
| FDGMM1                | 0.149         | FDGMM2                | 0.192          |
| Iter. Bayes           | 0.159         | Iter. Emp. Bayes      | 0.192          |
| RE(SWAR) with $Z_h$   | 0.160         | IOLS                  | 0.205          |
| FDGMM2                | 0.168         | POLS                  | 0.251          |
| RE(SWAR)              | 0.168         | RE(WALHUS)            | 0.259          |
| RE(WALHUS)            | 0.170         | RE(MLE)               | 0.262          |
| RE(MLE) with $Z_h$    | 0.172         | I2SLS                 | 0.285          |
| POLS                  | 0.172         | POLS with $Z_h$       | 0.319          |
| POLS with $Z_h$       | 0.179         | RE(MLE) with $Z_h$    | 0.330          |
| GMM2                  | 0.179         | Iter. Bayes           | 0.332          |
| RE(WALHUS) with $Z_h$ | 0.183         | GMM2                  | 0.340          |
| GMM1                  | 0.184         | FD2SLS-L              | 0.344          |
| FD2SLS                | 0.184         | RE(WALHUS) with $Z_h$ | 0.347          |
| IOLS                  | 0.197         | GMM1                  | 0.348          |
| I2SLS                 | 0.197         | RE(SWAR)              | 0.362          |
| GMM2 with $Z_h$       | 0.210         | RE(SWAR) with $Z_h$   | 0.413          |
| GMM1 with $Z_h$       | 0.225         | GMM1 with $Z_h$       | 0.424          |
| 2SLS with $Z_h$       | 0.225         | GMM2 with $Z_h$       | 0.437          |
| FD2SLS-L              | 0.225         | 2SLS with $Z_h$       | 0.504          |
| Emp. Bayes            | 0.366         | 2SLS                  | 0.628          |
| 2SLS                  | 0.418         | Emp. Bayes            | 0.724          |
| MG                    | 1.079         | MG                    | 1.131          |

RMSE = Root mean squared error.