ZAF (2011) 43:277–297
DOI 10.1007/s12651-010-0046-2

RESEARCH PAPER

Testing the neoclassical migration model: overall and age-group specific results for German regions

Timo Mitze · Janina Reinkowski

Received: 4 July 2010 / Accepted: 15 November 2010 / Published online: 24 December 2010

Abstract This paper tests the empirical validity of the neoclassical migration model in predicting German internal migration flows. We estimate static and dynamic migration functions for 97 Spatial Planning Regions between 1996 and 2006 using key labor market signals including income and unemployment differences among a broader set of explanatory variables. In addition to an aggregate specification we also estimate the model for age-group related subsamples. Our results give empirical support for the main transmission channels identified by the neoclassical framework – both at the aggregate level as well as for age-group specific estimates. Thereby, the impact of labor market signals is tested to be of greatest magnitude for workforce relevant age groups and especially young cohorts between the ages of 18 to 25 and 25 to 30. This latter result emphasizes the prominent role played by labor market conditions in determining internal migration rates of the working population in Germany.

Keywords German internal migration · Harris–Todaro model · Dynamic panel data

JEL classification R23 · C31 · C33

T. Mitze (✉)
RWI,
Hohenzollernstr. 1–3, 45128 Essen, Germany
e-mail: Mitze@rwi-essen.de
Tel.: +49-201-8149-223
Fax: +49-201-8149-200

T. Mitze
Ruhr University Bochum,
Bochum, Germany

J. Reinkowski
ifo Munich,
Munich, Germany

Tests zur Validität des neoklassischen Migrationsmodells: Allgemeine und altersgruppenspezifische Resultate für deutsche Raumordnungsregionen

Zusammenfassung Dieses Papier untersucht die empirische Validität des neoklassischen Migrationsmodells zur Erklärung interner Migrationsströme in Deutschland. Die Schätzungen basieren auf Daten für 97 Raumordnungsregionen im Zeitraum 1996 bis 2006. Neben einer aggregierten Migrationsgleichung werden auch disaggregierte Modelle für verschiedene Altersgruppen geschätzt. Sowohl die aggregierten als auch altersgruppenspezifischen Resultate bestätigen den durch das neoklassische Migrationsmodell vorhergesagten Zusammenhang zwischen interner Migration und regionalen Arbeitsmarktungleichgewichten. Dabei zeigen die Ergebnisse, dass insbesondere junge Alterskohorten zwischen 18 und 25 sowie 25 und 30 Jahren stark auf Unterschiede im regionalen Lohnniveau und der Arbeitslosenquote reagieren.

1 Introduction

There are many theories aiming to explain why certain people migrate and others do not. However, the neoclassical model remains the standard workhorse specification for analyzing internal and external migration rates at regional, national and international levels. The model places special emphasis on the labor market dimension of migration and basically relates migration-induced population changes to the relative income (or wage) and employment situation found in the regions of origin and destination.

In its response, migration works as an equilibrating mechanism for balancing differences among regions with respect to key labor market variables since higher in-migration in a region is expected to reduce the regional wage level due to an increase in labor supply. From the perspective of economic policy making, the empirical implications of the neoclassical migration model are important in order to assess whether labor mobility can act as an appropriate adjustment mechanism in integrated labor markets facing asymmetric shocks. Though the neoclassical migration model is widely used as a policy simulation and didactic tool, international empirical evidence so far has provided rather mixed results.

In this paper, we therefore aim to check the validity of the neoclassical migration model using a panel of 97 German regions for the period 1996–2006. We are especially interested in taking a closer look at the role played by time dynamic adjustment processes driving the internal migration patterns. We also aim to identify the role of additional factors as well as regional amenities in explaining migratory movements and key labor market signals. Finally, we focus on the heterogeneity of the adjustment processes taking place when migration flows are disaggregated by age groups.

The remainder of the paper is organized as follows: Section 2 sketches the theoretical foundations of the neoclassical migration model. Building on its theoretical underpinnings, Sect. 3 discusses the estimation approach with a special focus on dynamic panel data models. Section 4 then presents a selected literature review for empirical studies dealing with the determinants of internal migration flows. Section 5 describes the data used and displays stylized facts for German internal migration flows and regional labor market trends. Section 6 presents the empirical results for the total sample as well as for different age groups. Apart from an economic interpretation of the estimation coefficients obtained, we also carefully look at likely model misspecifications such as cross-sectional dependence in the error terms. Section 7 concludes the paper.

2 The neoclassical migration model

Given the complex nature of the decision making process faced by individuals, there is a large variety of theoretical models available to explain the actual migration outcome. These models may either be classified as micro or macroeconomic in nature. Given the scope of this paper, in the following we focus on the latter class, which particularly addresses the labor market dimension of migratory flows. However, as for many macro relationships, the neoclassical migration model is also grounded on solid microeconomic foundations. Its derivation starts from a lifetime expected income (utility) maximization approach as specified in the classical work on the human capital model of migration (see Sjaastad 1962). The human capital model in fact views the process of migration as an investment decision, where the returns to migration in terms of higher wages associated with a new job should exceed the costs involved in moving.

Relaxing the assumption that prospective migrants have perfect information about the wage rates and job availabilities among all potential locations involved in their decision making process, Todaro (1969) proposed a model framework where migrants discount wages by the probability of finding a job in alternative regions. Throughout the decision making process, each individual compares the expected (rather than observed) income level he would obtain if were to stay in his home region (i) with the expected income we would obtain in the alternative region (j) and further accounts for 'transportation costs' of moving from region i to j.

Harris and Todaro (1970) further formalize this idea. The authors set up a model where the expected income from staying in the region of residence Y_{ii}^E is a function of the wage rate or income in region i (Y_i) and the probability of being employed ($\text{Prob}(\text{EMP}_i)$). The latter in turn is assumed to be a function of the unemployment rate in region i (U_i) and a set of further economic and non-economic determinants (X_i). The same setup holds for region j accordingly. Taking costs of moving from region i to j into account (C_{ij}), the individual's decision will be in favor of moving to region j if

$$Y_{ii}^E < Y_{ij}^E - C_{ij}\,, \tag{1}$$

where $Y_{ii}^E = f(\text{Prob}(\text{EMP}_i),\ Y_i)$ and $Y_{ij}^E = f(\text{Prob}(\text{EMP}_j),\ Y_j)$. The potential migrant weights the proposed wage level in the home and target regions with the individual probability of finding employment. Using this information, we can set up a model for the regional net migration rate (NM_{ij}) defined as regional in-migration flows to i from j relative to outmigration flows from i to j (possibly normalized by the regional population level), which has the following general form:

$$\begin{aligned}\text{INM}_{ij} - \text{OUTM}_{ij} &= \text{NM}_{ij}\\ &= f(Y_i, Y_j, U_i, U_j, X_i, X_j, C_{ij})\,.\end{aligned} \tag{2}$$

With respect to the theoretically motivated signs of the explanatory variables, the model predicts that an increase in the home region wage rate (or, alternatively, the real income level) *ceteris paribus* leads to higher net migration inflows, while a wage rate increase in region j results in a decrease of the net migration rate. On the contrary, an increase in the unemployment rate in region i (j) has negative (positive) effects on the bilateral net migration from i to j. The costs of moving from i to j are typically expected to be an impediment to migration and are negatively correlated with net migration as:

$$\frac{\partial \mathrm{NM}_{ij}}{\partial Y_i} > 0\,; \quad \frac{\partial \mathrm{NM}_{ij}}{\partial Y_j} < 0\,; \quad \frac{\partial \mathrm{NM}_{ij}}{\partial U_i} > 0\,; \quad \frac{\partial \mathrm{NM}_{ij}}{\partial U j} < 0\,; \quad \frac{\partial \mathrm{NM}_{ij}}{\partial C_{ij}} < 0\,. \tag{3}$$

Core labor market variables may nevertheless not be sufficient to fully predict regional migration flows. We may extend the model by further driving forces of migration such as human capital, the regional competitiveness, housing prices, population density and environmental conditions, among others (see, e.g. Napolitano and Bonasia 2010, for an overview). For notational purposes, in the following we refer to the neoclassical migration model solely focusing on labor market conditions as the 'baseline' specification, while the 'augmented' specification also checks for regional amenities and further driving forces such as the regional skill level, population density and commuting flows as a substitute for migratory movements.

The likely impact of additional variables in the augmented neoclassical framework can be sketched as follows. Taking human capital as an example, it may be quite reasonable to relax the assumption of the Harris–Todaro model that an uneducated laborer has the same chance of getting a job as an educated laborer. Instead, the probability of finding a job is also a function of the (individual but also region specific) endowment with human capital (HK). The same logic holds for regional competitiveness (INTCOMP). Here, we expect that regions with a high competitiveness are better equipped to provide job opportunities than regions lagging behind (where regional competitiveness may, e.g. be proxied by the share of foreign turnover relative to total turnover in sectors with internationally tradable goods). For population density (POPDENS), we expect a positive impact of agglomeration forces on net flows through an increased possibility of finding a job, given the relevance of spillover effects, e.g. from a large pooled labor market. Thus, the probability of finding employment in region i in the augmented neoclassical migration model takes the following form:[1]

$$\begin{aligned} \mathrm{Prob}(\mathrm{EMP}_i) &= f[U_i, \mathrm{HK}_i, \mathrm{INTCOMP}_i, \mathrm{POPDENS}_i]\,, \\ \text{with:} \quad & \frac{\partial \mathrm{NM}_{ij}}{\partial \mathrm{HK}_i} > 0\,; \frac{\partial \mathrm{NM}_{ij}}{\partial \mathrm{INTCOMP}_i} > 0\,; \\ & \frac{\partial \mathrm{NM}_{ij}}{\partial \mathrm{POPDENS}_i} > 0\,. \end{aligned} \tag{4}$$

Moreover, we also carefully account for alternative adjustment mechanisms such as interregional net commuting flows to restore the inter-regional labor market equilibrium along with migratory movements. As Alecke and Untiedt (2001) point out, the theoretical as well as empirical literature with respect to interregional commuting (different from intraregional commuting) is rather scarce. According to Evers (1989), theoretical models of interregional commuting base the commuting decision on driving forces similar to those outlined in the migration framework. We thus expect that these flows are negatively correlated with net in-migration after controlling for common determinants such as regional income differences.

Finally, regional amenities are typically included as a proxy variable for (unobserved) specific climatic, ecological or socio-economic conditions in a certain region. According to the amenity approach regional differences in labor market signals then only exhibit an effect on migration after a critical threshold has been passed. Since in empirical terms it is often hard to operationalize amenity relevant factors, Greenwood et al. (1991) proposed to test the latter effect by the inclusion (macro)regional dummy variables in the empirical model. For the long run net migration equation, amenity-rich regions should have dummy coefficients greater than zero, indicating that those regions exhibit higher than average in-migration rates, as would be expected after controlling for regional labor market and macroeconomic differences.

3 Econometric specification

3.1 Functional form of the empirical migration equation

For empirical estimation of the neoclassical migration model we start from its baseline specification as, e.g., applied by Puhani (2001) and set up a model for the net migration rate as:

$$\left(\frac{\mathrm{NM}_{ij,t}}{\mathrm{POP}_{i,t-1}}\right) = A_{i,t}\left(\frac{U_{i,t-1}^{\alpha_1} Y_{i,t-1}^{\alpha_2}}{U_{j,t-1}^{\alpha_3} Y_{j,t-1}^{\alpha_4}}\right), \tag{5}$$

where net migration rate between i and j is defined as regional net balance NM for region i relative to the rest of the country j, POP is the region's i population level, t is the time dimension.[2] A is a (cross-section specific) constant term. In the empirical literature, a log-linear stochastic form of the migration model in Eq. (5) is typically chosen, where lower case variables denote logs and $nmr_{ij,t} = \log(\mathrm{NM}_{ij,t}/\mathrm{POP}_{i,t-1})$ as

$$\begin{aligned} nmr_{ij,t} = \alpha_0 &+ \alpha_1 y_{i,t-1} + \alpha_2 y_{j,t-1} \\ &+ \alpha_3 u_{i,t-1} + \alpha_4 u_{j,t-1} + \alpha_5 \mathbf{X} + e_{ij,t}\,, \end{aligned} \tag{6}$$

[1] The opposite effect on NM_{ij} holds for an increase in HK ↑, INTCOMP ↑ and POPDENS ↑ in region j.

[2] See e.g. Maza and Villaverde (2004) for a similar definition of the dependent variable.

where $e_{ij,t}$ is the model's error term. Taking into account that migration flows typically show a degree of persistence over time, we augment Eq. (6) by including one-period lagged values of net migration

$$\begin{aligned} nmr_{ij,t} = \beta_0 + \beta_1 nmr_{ij,t-1} + \beta_2 y_{i,t-1} + \beta_3 y_{j,t-1} \\ + \beta_4 u_{i,t-1} + \beta_5 u_{j,t-1} + \beta_6 \mathbf{X} + e_{ij,t} . \end{aligned} \tag{7}$$

The inclusion of a lagged dependent variable can be motivated by the existence of social networks in determining internal migration flows over time. Rainer and Siedler (2009), for example, find for German micro data that the presence of family and friends is indeed an important predictor for migration flows in terms of communication links, which may result in a gradual adjustment process over time for migration flows out of a particular origin to a destination region.

To account for the role played by timely adjustment processes in the endogenous variable, in the context of panel data models specific estimation techniques based on instrumental variables have to be applied. Besides the problem arising from a dynamic model specification, these techniques, in combination with an appropriate lag selection for the further explanatory variables, it may also help to minimize the fundamental endogeneity problem in this model setup, which arises from a two-way causality between internal migration and regional labor market variables. We give a detailed discussion of the latter point throughout the outline of the applied estimation techniques in the following.

Finally, in applied work one typically finds a restricted version of Eq. (7) where net migration is regressed against regional differences of explanatory variables of the form (see, e.g. Puhani 2001)

$$\begin{aligned} nmr_{ij,t} = \gamma_0 + \gamma_1 nmr_{ij,t-1} + \gamma_2 \tilde{y}_{ij,t-1} + \gamma_3 \tilde{u}_{ij,t-1} \\ + \gamma_4 \mathbf{X} + e_{ij,t} , \end{aligned} \tag{8}$$

where $\tilde{x}_{ij,t}$ for a variable $x_{ij,t}$ denotes $\tilde{x}_{ij,t} = x_{i,t} - x_{j,t}$. The latter specification implies the following testable restrictions

$$\beta_2 = -\beta_3 , \tag{9}$$

$$\beta_4 = -\beta_5 . \tag{10}$$

3.2 Choice of estimation technique and model misspecification tests

For estimation purposes we then have to find an appropriate estimator that is capable of handling the above described empirical setup. Given the dynamic nature of the neoclassical migration model in Eq. (7), we can write the specified form in terms of a more general dynamic panel data model as (in log-linear specification):

$$\begin{aligned} y_{i,t} = \alpha_0 + \alpha_1 y_{i,t-1} + \sum_{j=0}^{k} \beta_j' X_{i,t-j} + u_{i,t} , \\ \text{with:} \quad u_{i,t} = \mu_i + \nu_{i,t} , \end{aligned} \tag{11}$$

again $i = 1, \ldots, N$ (cross-sectional dimension) and $t = 1, \ldots, T$ (time dimension). $y_{i,t}$ is the endogenous variable and $y_{i,t-1}$ is one-period lagged value. X_i is the vector of explanatory time-varying and time invariant regressors, $u_{i,t}$ is the combined error term, where $u_{i,t}$ is composed of the two error components μ_i as the unobservable individual effects and $\nu_{i,t}$ is the remaining error term. Both μ_i and $\nu_{i,t}$ are assumed to be i.i.d. residuals with standard normality assumptions.

There are numerous contributions in the recent literature on how to estimate a dynamic model of the above type, which especially deal with the problem introduced by the inclusion of a lagged dependent variable in the estimation equation and its built-in correlation with the individual effect; that is, sincet y_{it} is a function of μ_i, also $y_{i,t-1}$ is a function of μ_i and thus $y_{i,t-1}$ as right-hand side regressor in Eq. (11) is likewise correlated with the combined error term. Even in the absence of serial correlation of ν_{it} this renders standard λ-class estimators such as OLS, the fixed effects model (FEM) and the random effects model (REM) inconsistent (see, e.g. Nickell 1981, Sevestre and Trogon 1995 or Baltagi 2008, for an overview).

Next to direct approaches aiming to correct for the bias of the FEM (see, e.g. Kiviet 1995, Everaert and Pozzi 2007, and the related literature for analytical or bootstrapping-based correction factors), the most widely applied approaches of dealing with this kind of endogeneity typically applies instrumental variable (IV) and generalized methods of moments (GMM) based techniques. While the first generation of models used transformations in first differences, latter extensions also account for the information in levels, when setting up proper estimators. A common tool is the system GMM estimator by Blundell and Bond (1998) as weighted average of first difference and level GMM.

Especially the latter estimators are a good candidate to simultaneously handle the problem arising from the inclusion of the lagged migration variable in our empirical model and the fundamental endogeneity problem induced by two-way causality between migration and labor market variables. In our case, the combination of an appropriate lag selection for the right-hand side regressors combined with the IV approach may do so. That is, since we include labor market variables with a lag structure in Eq. (7), by definition there cannot be any direct feedback effect from $nmr_{ij,t}$ to labor market variables. However, since $nmr_{ij,t-1}$ enters contemporaneously with respect to the latter, there is still the risk of two-way interdependencies due to the dynamic setting of the model. We minimize these potential risks of any endogeneity bias by instrumenting $nmr_{ij,t-1}$ with its lagged values so that the possibility of feedback effects from migration responses to labor market changes as source of estimation bias is limited. This should lead to consistent estimates of the coefficients for the explanatory variables.[3]

We are then also particularly interested in testing for the appropriateness of the chosen IV approach and apply test routines that account for the problem of many and/or weak instruments in the regression (see, e.g. Roodman 2009). Moreover, as it is typically the case with regional data, we are especially aware of the potential bias induced by a significant cross-sectional dependence in the error term of the model. There are different ways to account for such error cross-sectional dependences implying $\mathrm{Cov}(\nu_{i,t}\nu_{j,t}) \neq 0$ for some t and $i \neq j$ (see Sarafidis and Wansbeek 2010).

Besides the familiar spatial econometric approach, which assumes certain distance decay in spatial dependence, recently the common factor structure approach has gained considerable attention. The latter specification assumes that the disturbance term contains a finite number of unobserved factors that influence each individual cross-section separately. The common factor model approach is based on the concept of strong cross-sectional dependence, which assumes that all regions, either symmetrically or asymmetrically, are affected rather than just those nearby. Common examples are, for instance, regional adjustment processes to common macroeconomic shocks. We introduce a common factor structure for the error term according to Eq. (11) in the following way:

$$u_{i,t} = \mu_i + \nu_{i,t}\,, \quad \nu_{i,t} = \sum_{m=1}^{M} \phi_{m,i}\mathbf{f}_{m,t} + \epsilon_{i,t}\,, \tag{12}$$

where $\mathbf{f}_{m,t} = (f_{1,t}, \ldots, f_{M,t})'$ denotes an $M \times 1$ vector of individual-invariant time-specific unobserved effects, $\phi_i = (\phi_{1,i}, \ldots, \phi_{M,i})'$ is an $M \times 1$ vector of factor loadings and $\epsilon_{i,t}$ is a pure idiosyncratic error component with zero mean and constant variance. Cross-sectional dependence in turn leads to inconsistent estimates if regressors are correlated with the unspecified common variables or shocks. There are different proposals in the literature on how to account for unobserved factors.

For dynamic panel estimators with short time dimension, Sarafidis and Robertson (2009) propose applying time-specific demeaning, which alleviates the problem of parameter bias if the variance of the individual factor loadings for the common factor models is small. Alternatively, if the impact of the common factor varies considerably by cross-sections, there are different estimation techniques that account for this type of cross-sectional dependence by using cross-section averages of the dependent and independent variables as additional regressors (see, e.g. Pesaran 2006).

Recently, various testing procedures have been developed to check for the presence of cross-sectional dependence. Among the most commonly applied routines is Pesaran's (2007) extension to the standard Breusch and Pagan LM test. The so-called Cross-Section Dependence (CD) test is based on the pairwise correlation coefficient of residuals from a model specification that ignores the potential presence of cross-sectional dependence. However, as Sarafidis and Wansbeek (2010) point out, the CD-Test has the weakness that it may lack power to detect the alternative hypothesis under which the sign of the elements of the error covariance matrix alternates (thus for positive and negative correlation in the residuals, e.g. for factor models with zero mean factor loadings).

Moreover, the test statistic requires normality of the residuals. Sarafidis et al. (2009) propose an alternative testing procedure that does not require normality and is valid for fixed T and large N. The testing approach, which is designed for the Arellano and Bond (1991) and Blundell and Bond (1998) GMM estimators, is based on the Diff-in-Hansen test for overidentifying restrictions. The latter is also known as the C-statistic and is defined according to Eichenbaum et al. (1988) as the difference between two Sargan (1958)/Hansen (1982) J-statistics for an unrestricted and restricted IV/GMM-model. The aim of the test is to examine whether there is still (heterogeneous) cross-sectional dependence in the residuals after time-specific demeaning in the logic of Sarafidis and Robertson (2009). The test has the following form:

$$C_{\text{CD-GMM}} = (S_{\mathrm{F}} - S_{\mathrm{R}}) \xrightarrow{d} \chi^2_{h_{\mathrm{d}}}\,, \tag{13}$$

where h_{d} is the number of degrees of freedom of the test statistic as difference between the set of instruments (number of moment conditions) in the full model (S_{F}) and the restricted model (S_{R}), where the GMM model has either the Arellano–Bond or the Blundell–Bond form augmented by time-specific dummy variables. The corresponding null hypothesis of the Sargan's difference-test tests is that there is homogeneous cross-sectional dependence in the model versus the alternative of heterogeneous cross-sectional dependence. If only homogeneous cross-sectional dependence is present, the inclusion of time-specific dummies variables is sufficient to remove any bias in the estimation approach, see, e.g. Sarafidis and Robertson (2009).[4]

[3] Of course, a full account of the simultaneity problem may call for a system approach that is also likely to increase the estimation efficiency if there are significant cross-correlations in the error terms for functional forms of the migration and labor market variable equations. However, a fully specified system approach goes beyond the scope of this paper.

[4] The restricted (sub)set of moment conditions thereby only includes instruments from regressors in the vector $X_{i,t}$ (according to Eq. (11)) that

4 What does the empirical literature say?

Testing for the empirical validity of the neoclassical migration model yields rather mixed results, when looking at recent empirical evidence for European data. Here, regional (un)employment disparities are often shown to be important factors in determining migratory flows. On the contrary, the influence of regional wage or income levels is difficult to prove in many empirical examinations (see, e.g. Pissarides and McMaster 1990, as well as Jackman and Savouri 1992, for British regions; Westerlund 1997, for inter-regional migration in Sweden; Devillanova and Garcia-Fontes 2004, for Spain). For the Italian case, Daveri and Faini (1999) show that the regional wage level corresponds to the theoretically expected signal for the gross outward migration from southern to northern regions. Similar results are found in Fachin (2007).

Napolitano and Bonasia (2010) show that although the coefficients for Italian labor market variables in the neoclassical migration model show the expected sign, due to the complexity of the internal migration process, the baseline Harris–Todaro approach neglects important variables such as agglomeration forces measured by population density and human capital. The latter variables are also found to be significant in addition to the standard labor market variables in an inter-regional migration model for the Polish transition process (see Ghatak et al. 2008). This indicates that the augmented migration model may be in order.

Turning to the case of German interregional migration, Decressin (1994) examined gross migration flows for West German states up to 1988. His results show that a wage increase in one region relative to others causes a disproportional rise in the gross migration levels in the first region. On the other hand, a rise in the unemployment in a region relative to others disproportionally lowers the gross migration levels. Decressin does not find a significant connection between bilateral gross migration and regional differences in wage level or unemployment when purely cross-sectional estimates are considered.

Difficulties in proving a significant influence of regional wage decreases on the migratory behavior within Germany are also found in earlier empirical studies based on micro-data directly addressing the motivation for individual migratory behavior in Germany. Among these are Hatzius (1994) for the West German states, and Schwarze and Wagner (1992), Wagner (1992), Burda (1993) and Büchel and Schwarze (1994) for East Germany. Subsequent studies succeed in qualifying the theoretically unsatisfactory result of an insignificant wage influence. Schwarze (1996) shows that by using the expected wage variables instead of the actual ones, the wage drop between East German and West German states has a significant influence on the migratory behavior.[5] In a continuation of the work (Burda 1993), Burda et al. (1998) also indicate a significant non-linear influence on household income.

Contrary to earlier evidence, in recent macroeconomic studies with an explicit focus on intra-German East–West migration flows, regional wage rate differentials are broadly tested to significantly affect migration flows (see, e.g. Parikh and Van Leuvensteijn 2003, Burda and Hunt 2001, Hunt 2006, as well as Alecke et al. 2010). The study of Parikh and Van Leuvensteijn (2003) augments the core migration model with regional wage and unemployment differentials as driving forces of interregional migration by various indicators such as regional housing costs, geographical distance and inequality measures. For the sample period 1993 to 1995, the authors find a significant non-linear relationship between disaggregated regional wage rate differences and East–West migration (of a U-shaped form for white-collar workers and of inverted U-form for blue-collar workers), while unemployment differences showed to be insignificant. The relationship between income inequality and migration did not turn out to be strong.

According Burda and Hunt (2001), wage rate differentials and especially the fast East–West convergence are also a significant indicator in explaining the state-to-state migration patterns observed. Using data from 1991 to 1999, the authors find that the decline in East–West migration starting from 1992 onwards can almost exclusively be explained by wage differentials and the fast East–West wage convergence, while unemployment differences do not seem to play an important part in explaining actual migration trends. The study by Hunt (2006) comes closest to the research focus in this paper. Hunt (2006) also estimates the migration response to labor market signals by age groups. The author finds that young potential emigrants are more sensitive to wages than older cohorts. At the same time young age groups are found to be less sensitive to unemployment levels in the origin region. Hunt (2006) argues that the latter finding is likely to drive the migration pattern pooled over all age groups and thus gives a motivation for the dominance of wage rate signals in aggregate data as, e.g., reported by Burda and Hunt (2001).

Alecke et al. (2010) apply a Panel VAR to analyze the simultaneous impact of labor market variables to migration

remain strongly exogenous in the sense that their factor loadings are mutually uncorrelated with the cross-section specific parameter of the common factor. Sarafidis et al. (2009) propose to likewise test for the exogeneity of a subset of regressors by means of the standard Sargan/Hansen test for overidentifying restrictions in a first step.

[5] This result is also confirmed by Brücker and Trübswetter (2004). The latter study also focuses on the role of self-selection in East–West migration, finding that East–West migrants receive a higher individual wage compared to their non-migrating counterparts after controlling for the human capital level.

Table 1 Variable definition and data sources

Variable	Description	Source
NM	Net migration defined as in- minus outmigration	Destatis (2009)
NM(to18)	Net migration of persons under 18 years	Destatis (2009)
NM(18to25)	Net migration of persons aged between 18 and 24	Destatis (2009)
NM(25to30)	Net migration of persons aged between 25 and 29	Destatis (2009)
NM(30to50)	Net migration of persons aged between 30 and 49	Destatis (2009)
NM(50to65)	Net migration of persons aged between 50 and 65	Destatis (2009)
NM(over65)	Net migration of persons aged 65 and above	Destatis (2009)
POP	Population Level	VGRdL (2009)
Y	Gross Domestic Product (real) per Capita	VGRdL (2009)
UR	Unemployment Rate	Federal Employment Agency (2009)
COMM	Net Commuting level defined as in- minus out-commuting	Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR 2009)
HK	Human Capital level defined as %-share of employees with university degree relative to total employees	BBSR (2009)
INTCOMP	International Competitiveness proxied by foreign turnover relative to total turnover in manufacturing industries	BBSR (2009)
EAST	Binary dummy variable for regions in East Germany	own calculation
STATE	Set of binary dummies for each of the 16 Federal States	own calculation
TIME	Set of year specific time dummies for sample period 1996 to 2006	own calculation
SETTLE	Set of binary dummies for types of settlement structure with: *Type1*: Highly agglomerated area with regional urban center above 100.000 persons and population density above 300 inhabitants/sqm *Type2*: Highly agglomerated area with regional urban center above 100.000 persons and population density below 300 inhabitants/sqm *Type3*: Agglomerated area with population density above 200 inhabitants/sqm *Type4*: Agglomerated area with regional urban center above 100.000 persons and population density between 100–200 inhabitants/sqm *Type5*: Agglomerated area without regional urban center above 100.000 persons and population density between 150–200 inhabitants/sqm *Type6*: Rural area with population density above 100 inhabitants/sqm *Type7*: Rural area with population density below 100 inhabitants/sqm	BBSR (2009)
i	index for region i (region in focus)	
j	index for region j (rest of the country aggregate)	
t	time index	

and vice versa for German federal states between 1991 and 2006. The results broadly support the neoclassical migration model and show that migration itself has an equilibrating effect on labor market differences. The authors also find evidence for structural differences between the German west and east macro regions in the migration equation, which is similar to findings for an Italian ‘empirical puzzle’ with a distinct North–South division in terms of the magnitude of migration responses to labor market signals (see, e.g. Fachin 2007, and Etzo 2007).

The recent results for Germany also show that the specific time period used for estimation may have significant impact on the estimation results. Especially for the first years after the reunification several structural breaks are in order that may partly explain the results between earlier and recent contributions with respect to German internal migration. However, except for Alecke et al. (2010), none of the empirical papers take into account recent sample observations incorporating information about the second wave of strong East–West outmigration around the year 2001. The allocation of higher weights to recent sample observations may in turn minimize the risk of biasing the results in the light of distinct macro regional structural breaks.[6]

5 Data and stylized facts

We use the heterogeneous findings in the international and German empirical literature regarding the neoclassical mi-

[6] In this paper we account for regional and macro regional results by including East German and state level fixed effects. However, future work should also explicitly test for the poolability of the data for regional subgroups in a partial clustering framework.

Table 2 Descriptive statistics for continuous variables in the sample

Variable	Obs.	Mean	Std. dev.	Min	Max	Unit
INM	1067	0.00	7.21	−95.90	37.01	in 1000 persons
INM (to18)	1067	0.00	1.91	−24.41	32.41	in 1000 persons
INM (18to25)	1067	0.00	1.85	−12.97	15.76	in 1000 persons
INM (25to30)	1067	0.00	1.27	−9.93	12.42	in 1000 persons
INM (30to50)	1067	0.00	2.48	−30.99	8.24	in 1000 persons
INM (50to65)	1067	0.00	0.91	−10.61	1.82	in 1000 persons
INM (over65)	1067	0.00	0.62	−7.05	1.23	in 1000 persons
POP	1067	848.10	607.13	226.29	3466.52	in 1000 persons
Y	1067	51.23	7.49	34.02	80.01	in 1000 Euro
UR	1067	11.84	4.94	4.37	26.18	in %
COMM	873	−33.49	37.44	−177.73	36.31	in 1000 persons
HK	873	7.30	2.71	2.88	16.81	in %
INTCOMP	946	30.05	11.42	0.82	61.12	in %

gration model as a starting point for an updated regression approach based on German spatial planning units between 1996 and 2006. For empirical estimation we use regional data for the 97 German Spatial Planning Regions (so called *Raumordnungsregionen*) as the level of analysis for spatial migration processes within Germany (see, e.g. Bundesinstitut für Bau-, Stadt-, und Raumforschung 2010, for details about the concept of Spatial Planning Regions).[7]

We use a set of variables comprising regional net migration, population, real income, unemployment rate, human capital endowment, international competitiveness of regions and commuting flows. The latter have been included to account for an alternative adjustment mechanism to balance labor market disequilibria. Human capital is defined as the percentage share of regional employment of people with a university degree (including universities of applied science) in the total employment covered by the social security system (*sozialversicherungspflichtig Beschäftigte*).[8] We also include two sets of dummy variables. 1. Binary dummy variables for the 16 federal states to capture macro regional differences (see, e.g. Suedekum 2004). This may be especially important to account for structural differences between West and East Germany (see, e.g. Alecke et al. 2010, for recent findings). 2. Binary dummy variables for different regional settlement types ranging from metropolitan agglomerations to rural areas (in total 7 different categories based on their absolute population size and population density). As Napolitano and Bonasia (2010) point out, variables measuring population density may be an important factor in explaining the regional amenities. Variable definitions and descriptive statistics are provided in Tables 1 to 3.

Table 3 Descriptive statistics for binary variables in the sample

Variable	Obs.	% with $X = 1$
EAST	1067	23.7
Federal state level dummies		
BW	1067	12.4
BAY	1067	18.5
BER	1067	1.0
BRA	1067	5.2
BRE	1067	1.0
HH	1067	1.0
HES	1067	5.1
MV	1067	4.1
NIE	1067	13.4
NRW	1067	13.4
RHP	1067	5.1
SAAR	1067	1.0
SACH	1067	5.1
ST	1067	4.1
SH	1067	5.1
TH	1067	4.1
Settlement type dummies		
Type1	1067	15.5
Type2	1067	15.5
Type3	1067	17.5
Type4	1067	17.5
Type5	1067	8.2
Type6	1067	15.4
Type7	1067	10.3

Note: BW = Baden-Württemberg, BAY = Bavaria, BER = Berlin, BRA = Brandenburg, BRE = Bremen, HH = Hamburg, HES = Hessen, MV = Mecklenburg-Vorpommern, NIE = Lower Saxony, NRW = North Rhine-Westphalia, RHP = Rhineland-Palatine, SAAR = Saarland, SACH = Saxony, ST = Saxony-Anhalt, SH = Schleswig-Holstein, TH = Thuringia.

[7] We restrict our estimation approach to this period since regional boundaries of the German Spatial Planning Regions changed before and after, which may introduce a measurement problem that is likely to bias our empirical results.

[8] We also checked for the sensitivity of the results when using composite indicators of human capital as discussed by Dreger et al. (2009), accounting for human capital potential (measured in terms of high school graduates with university qualification per total population between 18–20 years) as well as science and technology related indicators (e.g., patent intensity). The results did not change.

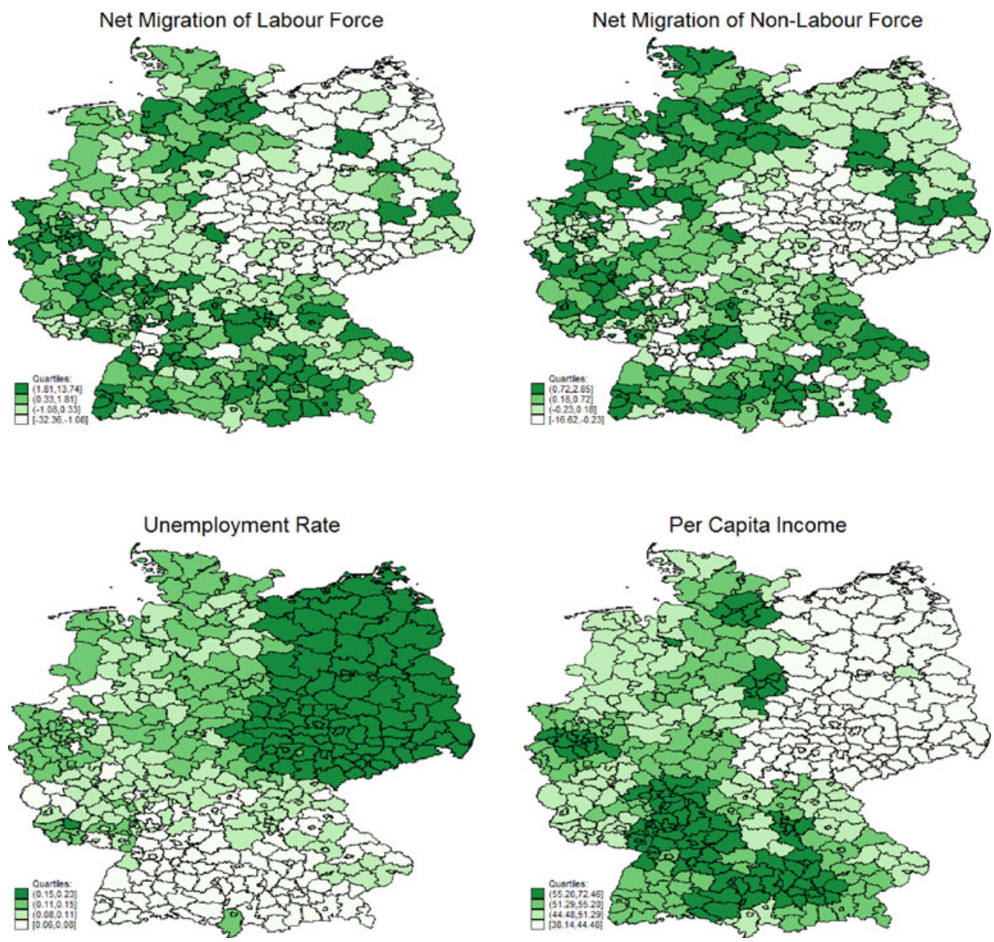


Fig. 1 Sample means of net migration (in 1000). Unemployment rate (in %), per capita GDP (in 1000€)

To highlight regional and macro regional differences for net migration and explanatory variables, Fig. 1 visualizes spatial differences for the sample means of net in-migration and labor market variables for the period 1996–2006. Net in-migration flows are categorized into labor force relevant age groups between 18 and 65 years as well as non-labor force relevant age groups. For labor force migration, the figure shows that throughout the sample period the East German regions on average lost a considerable fraction of their population levels through net out-migration. Exceptions are the economic core regions around Berlin/Brandenburg and in the south-west of Saxony. Also, the western regions along the border to East Germany experienced net outflows. On the other hand, the northern West German regions around the urban agglomerations Hamburg and Bremen are among the net recipient regions as well as the western agglomerated regions in the Rhineland (around the metropolitan areas Cologne and Düsseldorf) and the southern West German regions in Baden Württemberg and Bavaria.

Looking at net migration trends for non-labor market relevant age groups, the picture is less clear. We can see from Fig. 1 that both the north German coastal regions as well as the southern border regions gain considerable population through net in-migration. This trend may be interpreted in terms of regional amenities such as topographical advantages, which attract migration flows. The relative difference is especially observable for the East German coastal zone in Mecklenburg-Vorpommern. The spatial distribution of real per capita income and unemployment rates nevertheless show a distinct West–East division. The regions with the highest income levels for the sample period are the northern regions around Hamburg, the Western regions in the Rhineland as well as large parts of the southern states Baden-Württemberg and Bavaria. Since these regions were also found to have large net in-migration flows (both overall as well as for the workforce relevant age groups), this may give a first hint of the positive correlation of migration flows and regional income levels as suggested by the neoclassical migration model. The opposite case is supposed to hold for large regional unemployment rates. Especially for the East German Spatial Planning Regions high unemployment rates seem to match with net population losses. To check the correlation of these variables more in depth, the next section presents the results of the estimation exercise.

Table 4 Results of panel unit root tests (p-values) for variables in the migration model

Test used:	p-val. LLC	Lags	p-val. IPS	Lags	p-val. CADF	Lags
	H_0: All series are non-stationary					
$nm_{ij,t}$	(0.00)	1.47	(0.03)	1.47	(0.00)	1.00
$u_{i,t}$	(0.00)	3.20	(0.00)	3.20	(0.00)	1.00
$u_{j,t}$	(0.99)	3.81	(0.00)	0.22	(0.00)	1.00
$y_{i,t}$	(0.00)	1.35	(0.00)	1.35	(0.00)	1.00
$y_{j,t}$	(0.00)	0.00	(0.00)	0.00	(0.00)	1.00
$\tilde{u}_{ij,t}$	(0.00)	3.30	(0.00)	3.30	(0.00)	1.00
$\tilde{y}_{ij,t}$	(0.00)	1.44	(0.00)	1.44	(0.00)	1.00

Note: LLC denotes the test proposed by Levin et al. (2003), IPS is the Im et al. (2003) test, CADF is the test proposed by Pesaran (2007). All unit root tests include a constant term; optimal lag length selected according to the AIC information criterion for the LLC and IPS test. The Pesaran CADF test includes one lag and a potential time trend in the estimation equation.

6 Empirical results for the neoclassical migration model

6.1 Aggregate findings

For the migration model of Eqs. (7) and (8) we apply different static and dynamic panel data estimators. Before estimating the empirical migration model we check the time series properties of the variables involved in order to avoid the risk of running a spurious regression for non-stationary variables (with moderate $T = 11$). We therefore report test results of different panel unit root tests including recently proposed methods by Levin et al. (2003) and Im et al. (2003), as well as Pesaran's (2007) CADF test. The latter approach has the advantage that it is relatively robust with respect to cross-sectional dependence in the variable, even if the autoregressive parameter is high (see, e.g. Baltagi et al. 2007, as well as de Silva et al. 2009, for extensive Monte Carlo simulation evidence). As the results in Table 4 show, for almost exclusively all variables and test specifications the null hypothesis of non-stationarity of the series under observation can be rejected.[9] Given this overall picture of the panel unit root tests together with the theoretically motivated assumption that migration flows are transitory processes between two labor market equilibria, it seems reasonable to handle the variables as stationary processes so that we can run untransformed regressions without running the risk of spurious regression results.

For estimation we start from an unrestricted presentation of the baseline model including the core labor market variables real income (y) and unemployment rates (u) and test for parameter constraints according to Eqs. (9) and (10). As the results in Table 5 show, for almost all model specifications the null hypothesis for equal parameter size cannot be rejected on the basis of standard Wald tests. Also, compared to the static specification in column 2, the relative root mean squared error (RMSE) criterion of the model strongly increases if we add a dynamic component to the migration equation. The relative RSME for each estimator is thereby computed as the ratio of the model's RMSE and the static POLS benchmark specification in column 1. A value smaller than one indicate that the model has a better predictive performance than the benchmark POLS.

As discussed above the λ-class estimators are potentially biased in a dynamic specification. Since the coefficient of the lagged dependent variable turns out to be highly significant, we also compute a bias-corrected FEM specification as well as the Arellano and Bond (1991) and Blundell and Bond (1998) system GMM estimators. According to the relative RMSE criterion the Blundell–Bond system GMM specification has the smallest prediction error. The coefficients for labor market signals are statistically significant and of the expected signs. Moreover, the SYS-GMM specification passes standard tests for autocorrelation in the residuals (m_1 and m_2 statistics proposed by Arellano and Bond 1991) as well as the Hansen J-statistic for instrument validity. The reported C-statistic for the exogeneity of the instruments in the level equation shows the validity of the augmented approach in extension to the standard Arellano–Bond first difference model.

We then use the SYS-GMM approach to test for the significance of different extensions of the baseline Harris–Todaro model. We start by including a dummy variable for the East German Spatial Planning Regions (see Table 6). The motivation for this approach is to test for the significance of the so-called East German empirical puzzle, where a relatively high degree of migratory interregional immobility was found to coexist with large regional labor market disparities. Fachin (2007) and Etzo (2007) report similar results for Italian South–North migration trends, while Alecke and Untiedt (2000) as well as Alecke et al. (2010)

[9] It was only for the (rest of the country) aggregate of the unemployment rate that the Levin–Lin–Chu test could not reject the null of non-stationarity. However, the LLC-test rejects the null hypothesis of an integrated time series if the unemployment rate is transformed into regional differences ($\tilde{u}_{ij,t}$).

Table 5 Baseline specifications of the neoclassical migration model for German spatial planning regions

Dep. var.: $nm_{ij,t}$	POLS	POLS	POLS	REM	FEM	FEMc	AB-GMM	SYS-GMM
$nm_{ij,t-1}$	–	–	0.90***	0.90***	0.78***	0.92***	0.84***	0.88***
	–	–	(0.011)	(0.011)	(0.022)	(0.031)	(0.001)	(0.001)
$u_{i,t-1}$	−0.74***	–	–	–	–	–	–	–
	(0.114)	–	–	–	–	–	–	–
$u_{j,t-1}$	0.64*	–	–	–	–	–	–	–
	(0.399)	–	–	–	–	–	–	–
$\tilde{u}_{ij,t-1}$	–	−0.72***	−0.05	−0.05	−0.32**	−0.28*	−0.53***	−0.19***
	–	(0.114)	(0.041)	(0.041)	(0.166)	(0.166)	(0.023)	(0.006)
$y_{i,t-1}$	0.07	–	–	–	–	–	–	–
	(0.315)	–	–	–	–	–	–	–
$y_{j,t-1}$	−0.14	–	–	–	–	–	–	–
	(0.378)	–	–	–	–	–	–	–
$\tilde{y}_{ij,t-1}$	–	0.07	0.12	0.12	−0.26	−0.10	0.25***	0.03**
	–	(0.314)	(0.108)	(0.112)	(0.372)	(0.374)	(0.066)	(0.014)
No. of obs.	1067	1067	1067	1067	1067	1067	1067	1067
No. of groups	97	97	97	97	97	97	97	97
No. of years	11	11	11	11	11	11	11	11
$\beta_{u_i} = -\beta_{u_j}$	(0.83)	–	(0.60)	(0.42)	(0.11)	(0.19)	(0.00)	(0.14)
$\beta_{y_i} = -\beta_{y_j}$	(0.76)	–	(0.60)	(0.24)	(0.39)	(0.59)	(0.58)	(0.14)
m_1 and m_2	–	–	–	–	–	–	(0.42)/(0.24)	(0.35)/(0.24)
J-stat. overall	–	–	–	–	–	–	Passed	Passed
C-stat. LEV-EQ	–	–	–	–	–	–	–	Passed
Time dummies (11)	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Relative RMSE	1	1.07	0.38	0.38	0.41	0.39	0.43	0.38

Note: ***, **, * = denote significance levels at the 1, 5 and 10% level, respectively. Standard Errors in brackets.

Table 6 Augmented neoclassical migration model for German spatial planning regions

$nm_{ij,t}$	SYS-GMM					
$nm_{ij,t-1}$	0.87***	0.87***	0.89***	0.87***	0.86***	0.89***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
$\tilde{u}_{ij,t-1}$	−0.33***	−0.52***	−0.25***	−0.58***	−0.86***	−0.86***
	(0.008)	(0.022)	(0.030)	(0.034)	(0.060)	(0.058)
$\tilde{y}_{ij,t-1}$	0.47***	0.48***	0.30***	1.25***	0.84***	1.05***
	(0.046)	(0.11)	(0.047)	(0.118)	(0.172)	(0.225)
EAST	0.29***	–	–	0.63***	–	–
	(0.016)	–	–	(0.045)	–	–
COMM	–	–	−0.02***	−0.02***	−0.05***	−0.05***
	–	–	(0.002)	(0.002)	(0.006)	(0.007)
HK	–	–	–	–	–	0.004
	–	–	–	–	–	(0.011)
INTCOMP	–	–	–	–	–	0.05**
	–	–	–	–	–	(0.021)
	Type of settlement structure					
Type 2	–	–	–	−0.07**	−0.53***	−0.40***
	–	–	–	(0.035)	(0.143)	(0.126)
Type 3	–	–	–	0.01	−0.10	−0.02
	–	–	–	(0.039)	(0.083)	(0.088)
Type 4	–	–	–	−0.12***	−0.24***	−0.16*
	–	–	–	(0.041)	(0.085)	(0.082)
Type 5	–	–	–	0.02	−0.12	−0.01
	–	–	–	(0.049)	(0.088)	(0.095)
Type 6	–	–	–	−0.05	−0.08	0.04
	–	–	–	(0.047)	(0.094)	(0.107)
Type 7	–	–	–	−0.05	−0.29***	−0.15
	–	–	–	(0.045)	(0.110)	(0.117)
No. of obs.	1067	1067	873	873	873	753
Time dummies (11)	167.9***	12.4***	32.3***	12.8***	16.5***	6.4***
State dummies (16)	No	21.7***	No	No	26.6***	27.8***
m_1	(0.38)	(0.37)	(0.50)	(0.57)	(0.55)	(0.64)
m_2	(0.24)	(0.24)	(0.21)	(0.20)	(0.20)	(0.20)
J-stat. overall	(0.52)	(0.67)	(0.16)	(0.12)	(0.31)	(0.22)
C-stat. LEV-EQ	(0.99)	(0.99)	(0.76)	(0.63)	(0.97)	(0.57)
C-stat. exog. var.	(0.07)	(0.99)	(0.00)	(0.00)	(0.33)	(0.11)
C-stat. CD-GMM	–	(0.58)	–	–	(0.35)	(0.57)

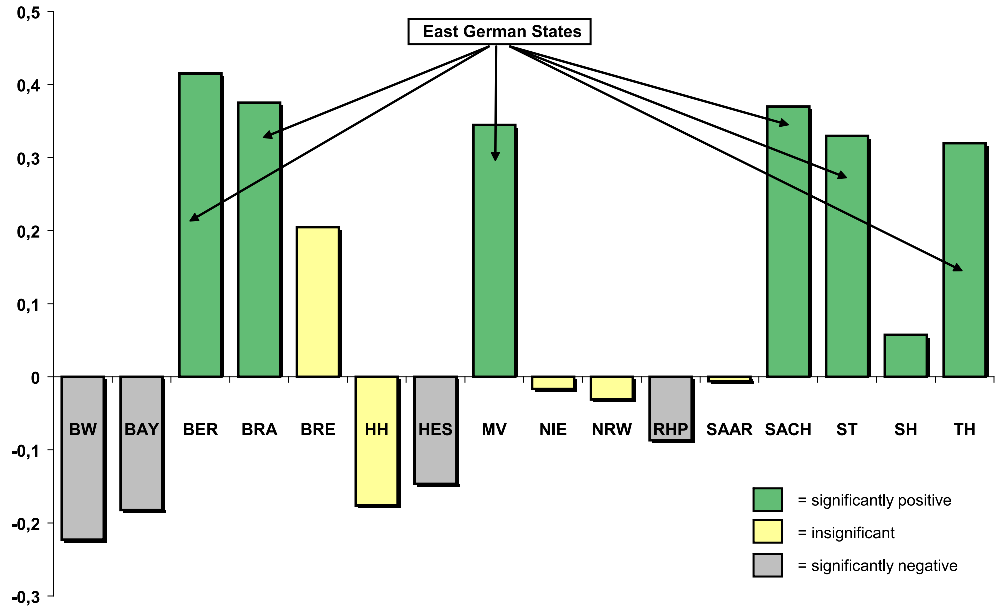
Note: ***, **, * = denote significance levels at the 1, 5 and 10% level, respectively. In the regressions including the regional settlement structure the dummy for highly agglomerated areas of Type 1 is excluded and thus serves as the benchmark category for the further settlement type dummies. Standard Errors in brackets. For m_1, m_2, J- and C-statistic test results p-values are reported.

identify such effects for German East–West migration throughout the 1990s.[10]

The results in Table 6 for the period 1996 to 2006 report a statistically significant positive East German dummy that indicates higher net in-migration balances for the East German spatial planning regions than their labor market performance would suggest. To obtain further insights we also estimate a specification that includes federal state level fixed effects. The estimation results for the state dummies in the baseline model are shown in Fig. 2. As the figure highlights, for all six East German state dummies we obtain statistically significant and positive coefficients. Negative coefficients are found for the West German states Baden Württemberg, Bavaria and Hessen. A Wald test for the joint effect of the set of state dummies turns out to be highly significant. However, most importantly, for both models including the East German dummy and the set of state dummies, the impact of labor market variables is still of the expected sign and higher than in the baseline specification. In line with Suedekum (2004) for West Germany, the results thus show that macro regional differences matter, yet there are no qual-

[10] The latter study found that along with a second wave of East–West movements in early 2000 net flows out of East Germany were much higher than expected after controlling its labor market and macroeconomic performance. Since this trend was accompanied by a gradual fading out of economic distortions, this supports the view of ‘repressed’ migration flows for that period.

Fig. 2 State level effects for German states in the aggregate baseline migration model



itative effects on the estimated coefficients that hint to a systematic rejection of the neoclassical migration model.

Regarding further variables in the augmented variable set, the results show that higher interregional net in-commuting levels are negatively correlated with the net in-migration rate. This supports our basic theoretical expectations above that both types are alternative adjustment mechanisms to reduce labor market disparities. The binary dummy variables for different settlement types (classified by size of local urban centers and population density, see Table 1 for details) reveal further structural differences in inter-regional migration patterns. Next to rural areas with low population density, agglomeration regions of Types 2 and 4 also show significantly lower net in-migration rates relative to benchmark category Type 1 (highly agglomerated area with a regional urban center above 100,000 persons and population density above 300 inhabitants/sqm). This may hint at the role played by regional centers of agglomeration in attracting migration flows and may be interpreted in favor of a 're-urbanization' process in Germany for the period 1996 to 2006. Similar trends have also been reported by Swiaczny et al. (2008).[11]

Finally, testing for the effects of regional human capital endowments and international competitiveness shows mixed results. While the proxy for the latter variable in terms of foreign turnover relative to total turnover in manufacturing sector industries shows the expected positive effect on net in-migration, the regional endowment with human capital is found to be insignificant. This finding corresponds to recent results for Spain between 1995–2002, where regional differences in human capital were not found to be helpful in predicting internal migration flows (see Maza and Villaverde 2004). The latter may be explained by the fact that not the region's specific stock of human capital but rather the individual endowment of the prospective migrant is the appropriate level of measurement. However, the latter variable is not observable for regional data.

In order to check the appropriateness of our augmented SYS-GMM specifications, we perform a variety of postestimation tests for instrument appropriateness and temporal and cross-sectional dependence of the error term. The test results are reported in Table 6. With respect to IV appropriateness and temporal autocorrelation of the error terms, all model specifications show satisfactory results. In order to control cross-sectional error dependence due to unobserved common factors, we first add year dummies to our model specification, which also turn out to be jointly significant. We then apply Sargan's difference test for the SYS-GMM model ($C_{\text{CD-GMM}}$) as described above, in order to check for the nature of the cross-sectional dependence given the impact of unobserved common factors.

In order to run the test, we first need to judge whether the set of explanatory variables (excluding instruments for the lagged endogenous variable) is exogenous with respect to the combined error term. This can easily be tested by means of a Sargan/Hansen J-statistic based overidentification test. As the results in Table 6 show, only those model specification that include fixed state effects pass the overidentification test for the vector of explanatory variables. For these

[11] The authors argue that throughout the process of demographic change in Germany city core regions may gain in demographic terms from young migrants, while suburban and rural areas are expected to face increasing migration losses.

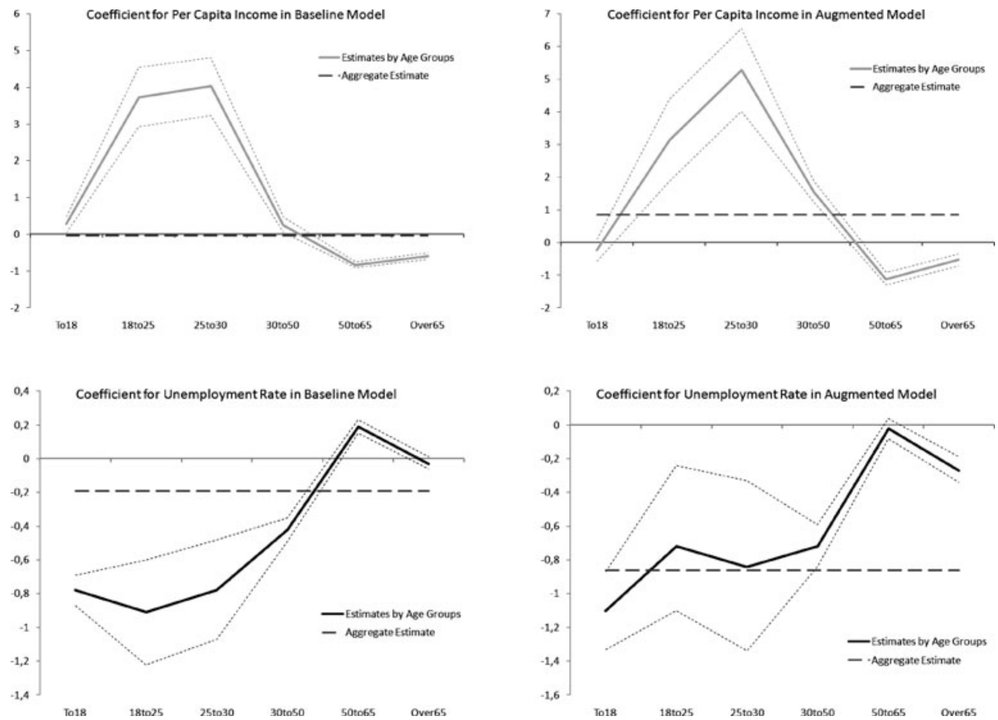


Fig. 3 Coefficients for income ($\tilde{y}_{ij,t-1}$) und unemployment rate differences ($\tilde{u}_{ij,t-1}$) by age groups

equations we can then apply $C_{\text{CD-GMM}}$ from Eq. (13) in order to test for the existence of heterogeneous factor loadings for the common factor structure of the error terms as proposed by Sarafidis et al. (2009). The test results do not indicate any sign of misspecification, including period-fixed effects for standard significance levels, hinting at homogeneous responses to common shocks. In sum, the augmented neoclassical migration equation is shown to be an appropriate representation of the data generating process and highlights the role of key labor market variables in explaining net in-migration rates in German regions.

6.2 Disaggregate estimates by age groups

Given the supportive findings for the neoclassical migration model at the aggregate level, we finally aim to check the sensitivity of the results when different disaggregated age groups are used. We are especially interested in analyzing whether the estimated coefficients for the labor market signals change for different age groups. Indeed, the estimation results show that the migratory response to labor market variables is much higher for workforce relevant age groups. For both the baseline and augmented model, the resulting coefficients for real income and unemployment rate differences together with 95% confidence intervals are plotted in Fig. 3.[12]

The coefficient for real income differences in Fig. 3 shows a clear inverted U-shaped pattern when plotted for the different age groups in ascending order. While for migrants aged up to 18 years real income differences do not seem to matter, for migrants aged between 18 to 25 and 25 to 30 years the estimated coefficient is statistically significant and much higher compared to the overall migration equation in Table 6. For older age groups the effect reduces gradually. The migration responses are found to be very similar for the baseline and augmented migration specification (see Fig. 3). Similar results were found for regional unemployment rate differences, which are shown to be almost equally important for age groups up to 50. It is only for elderly age groups that the coefficients turn out to be of smaller size and partly insignificant. If we look at the distribution of the state-level fixed effects for each estimated age-group specification, the estimation results show that the positive dummy variable coefficients for the East German states particularly hold for the workforce relevant age groups. The results are graphically shown in Fig. 4 for the baseline migration model.

Finally, Table 7 computes the ‘relative importance’ of the labor market variables by age groups in determining net migration flows. Thereby, the relative importance refers to the quantification of an individual regressor’s contribution in a multiple regression model (see, e.g. Grömping 2006, for an overview). This allows us to further answer the question as to how far our estimation results support the prominent role of labor market conditions in guiding internal migra-

[12] Detailed estimation results for the models are given in the Appendix.

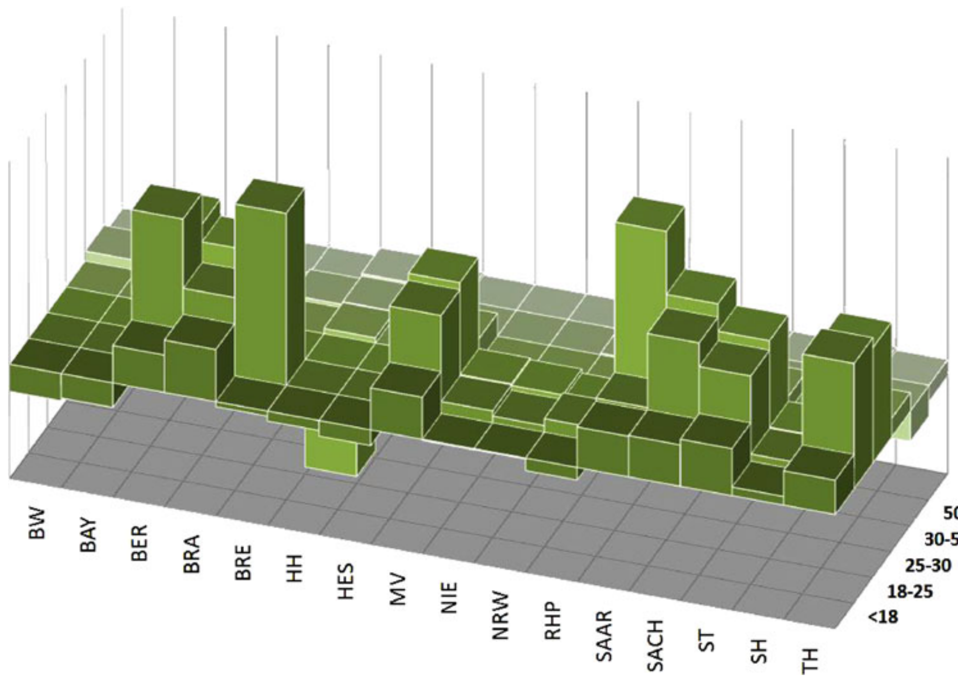


Fig. 4 State level effects in baseline migration model by states and age (Note: For details of calculation see Tables 8 and 9.)

Table 7 Relative contribution of labor market variables in explaining migration flows

Age group	Specification A $y_{ij,t-1}$	Specification A $u_{ij,t-1}$	Joint	Specification B $y_{ij,t-1}$	Specification B $u_{ij,t-1}$	Joint
Up to 18	1%	3%	4%	0%	19%	19%
18 to 25	29%	21%	50%	19%	8%	27%
25 to 30	18%	14%	31%	54%	11%	65%
30 to 50	1%	5%	6%	5%	8%	13%
50 to 65	1%	1%	1%	2%	0%	2%
Over 65	1%	0%	2%	1%	1%	2%

Note: Specification A is based on the computation of the squared correlation of the respective regressor with the dependent variables (univariate R^2). Specification B is calculated using the estimated SYS-GMM coefficient from the augmented migration model specification in Table 9 (see the Appendix). The estimation coefficient for regressor x_k is further standardized as $\hat{\beta}_{\text{standardized},k} = \hat{\beta}_k \sqrt{s_{kk}}/\sqrt{s_{yy}}$, where s_{kk} and s_{yy} denote the empirical variances of regressor x_k and the dependent variable y, respectively. As long as one only compares regressors within models for the same y, division by $\sqrt{s_{yy}}$ is irrelevant.

tion rates (of the workforce population) in Germany. Table 7 computes two specifications based on the squared correlation of the respective regressor with the dependent variables (univariate R^2, specification A) as well as the standardized estimated SYS-GMM coefficients from the augmented migration model. This latter metric for assessing the relative importance of regressors has an advantage over the simple benchmark in specification A since it accounts for the correlation of regressors. As the table shows, both methods assign a significant explanatory share to the two key labor market variables in predicting migration flows, especially for the workforce population (up to 50% joint contribution in Specification A for age group 18 to 25 years and even up to 65% for age group 25 to 30 in Specification B). The SYS-GMM thereby on average assigns a stronger weight to real income differences in explaining net in-migration relative to unemployment differences. However, the overall picture confirms our interpretation of the regression tables in assigning a prominent role to labor market imbalances in driving German internal migration.

7 Conclusion

In this paper, we have analyzed the explanatory power of the neoclassical migration model to describe aggregate and age-group specific internal migration trends for 97 German Spatial Planning regions throughout the period 1996–2006. Our results are based on model specifications for dynamic panel data estimators and give strong evidence in favor of the neoclassical inspired Harris–Todaro model. Both real income differences as well as unemployment rate disparities are found to be statistically significant with the expected signs. That is, a real income increase in region i relative to region j leads to higher net migration inflows to i from j; on the contrary, a rise in the regional unemployment rate in i leads to lower net inflows. Given these responses to labor market signals, migration flows may be seen as a spatial adjustment mechanism and equilibrate regional labor market imbalances.

The results of the standard neoclassical migration model remain stable if commuting flows, regional human capital endowment, the region's international competitiveness as well as differences in the settlement structure are added as further explanatory variables. The inclusion of the regional net in-commuting rate shows a negative correlation with migration underlying the substitutive nature of the two variables. Also, an increasing level of international competitiveness attracts further in-migration flows. We also find heterogeneity for different types of regional settlement structure proxied by population density and we observe persistent structural differences for the two East–West macro regions (by including individual federal state level fixed effects or a combined East German dummy). Most importantly, the impact of core labor market variables is still of the expected sign when further variables are added. In line with earlier empirical studies, the results thus show that macro regional differences matter, yet there are no qualitative effects on the estimated coefficients that hint to a systematic rejection of the neoclassical migration model.

We finally estimate the migration model for age-group specific subsamples of the data. Here, the impact of labor market signals is found to be of greatest magnitude for workforce relevant age groups (18 to 25, 25 to 30 and 30 to 50 years). Computing the 'relative importance' of labor market variables by age groups in a multiple regression framework with a broader set of controls, our results show that for young cohorts up to 65% of all migratory movements can be explained by differences in regional income levels and unemployment rates. This latter result emphasizes the prominent role played by labor market conditions in guiding internal migration rates of the working age population in Germany.

Executive summary

Given that the degree of geographical mobility of the labor force is an important indicator of the conduct of labor market policies, this paper tests the empirical validity of the neoclassical migration model in predicting German internal migration flows. The neoclassical migration model places special emphasis on the labor market dimension of migration and basically relates migration-induced population changes to the relative income (or wage) and employment situation found in the regions of origin and destination, respectively. Thereby, regions with a higher relative income level and lower unemployment rates are expected to exhibit positive net in-migration rates. In its response to regional labor market disparities, migration itself should work as an equilibrating mechanism for balancing these labor market differences. Higher in-migration flows in a certain region are expected to reduce the regional wage level due to an increase in labor supply and at the same time increase the regional unemployment rate. Thus, from the perspective of economic policy making, the empirical implications of the neoclassical migration model are important, for instance, to assess whether labor mobility can act as an appropriate adjustment mechanism in integrated labor markets facing asymmetric shocks. Though the neoclassical migration model is widely used as a policy simulation and didactic tool, the international empirical evidence so far has provided rather mixed results. To shed more light on the German case, we estimate static and dynamic migration functions for 97 Spatial Planning Regions between 1996 and 2006 using key labor market signals, including income and unemployment differences among a broader set of explanatory variables, including human capital, regional competitiveness, regional settlement structure and interregional commuting as alternative adjustment process for labor market disequilibria. In addition to an aggregate specification, we also estimate the model for age-group related subsamples. We estimate the migration equation by means of dynamic panel data estimation using instrumental variable techniques and place special emphasis on postestimation testing in order to control for right-hand side endogeneity and cross-sectional dependence of the error term.

Our results give empirical support for the main transmission channels identified by the neoclassical framework, both at the aggregate level as well as for age-group specific estimates. At the aggregate level, real income differences and unemployment rate disparities are found to be statistically significant with the expected signs. The results remain stable if commuting flows, regional human capital endowments, the region's international competitiveness as well as differences in the settlement structure are added as further explanatory variables. The inclusion of the regional net in-commuting rate shows a negative correlation with interregional migration, underlying the substitutive nature of the two variables. Further, an increasing level of international competitiveness attracts additional in-migration flows. We also find heterogeneous migration patterns for different types of regional settlement structure proxied by population density and we observe persistent structural differences for the two East–West macro regions (by including individual federal state level fixed effects or a combined East German dummy). However, it is important to note that the impact of core labor market variables is still of the expected sign when further variables are added. In line with earlier empirical studies, the results show that macro regional differences matter, yet there are no qualitative effects on the estimated coefficients

that hint toward a systematic rejection of the neoclassical migration model. Looking at the estimation results according to age groups, the impact of labor market signals is tested to be of greatest magnitude for workforce relevant age-groups and especially young cohorts between the ages of 18 to 25 and 25 to 30. We finally compute measures for the relative importance of labor market variables by age groups in a multiple regression framework including a broader set of controls. Our results show that for young cohorts up to 65% of the migratory movements can be explained by differences in regional income levels and unemployment rates. This latter result emphasizes the prominent role played by labor market conditions in determining internal migration rates of the workforce population in Germany.

Kurzfassung

Vor dem Hintergrund der arbeitsmarktpolitischen Relevanz von geografischen Mobilitätsraten der Erwerbsbevölkerung untersucht dieses Papier die empirische Validität des neoklassischen Migrationsmodells zur Erklärung interner Migrationsströme in Deutschland. Das neoklassische Migrationsmodell analysiert insbesondere die Arbeitsmarktdimension von Migrationsströmen und verbindet migrationsinduzierte Bevölkerungsveränderungen mit relativen Einkommens- und Beschäftigungssituationen in der Ursprungs- und Zielregion. Dabei wird angenommen, dass Regionen mit einem relativ höheren Einkommensniveau und niedriger Arbeitslosenquote tendenziell Migrationszuwächse verbuchen. Die induzierten Migrationsströme wirken in ihrer Reaktion auf regionale Arbeitsmarkungleichgewichte dann als Gleichgewichtsmechanismus, wobei eine gesteigerte In-Migration durch ein höheres Arbeitskräfteangebot den durchschnittlichen Lohnsatz in der Region reduziert und gleichzeitig die regionale Arbeitslosenquote anhebt. Aus wirtschaftspolitischer Perspektive sind die Implikationen des neoklassischen Migrationsmodells vor dem Hintergrund relevant, als dass so beispielsweise die Fähigkeit von Migrationsströmen zur Ausbalancierung asymmetrischer Schocks in regionalen Arbeitsmärkten untersucht werden kann. Obwohl das neoklassische Migrationsmodell im politischen Diskurs gängigerweise als Simulationsmodell und didaktisches Instrument eine breite Anwendung findet, ist die internationale empirische Evidenz hinsichtlich der zentralen Modellaussagen eher durchwachsen. Um mehr über die empirische Relevanz des neoklassischen Migrationsmodells für Deutschland zu erfahren, schätzt die Arbeit daher dynamische Migrationsgleichungen für 97 deutsche Raumordnungsregionen. Die Schätzungen basieren auf Daten aus dem Zeitraum von 1996 bis 2006 und beinhalten neben den zentralen Arbeitsmarktvariablen zur Messung regionaler Unterschiede in der Einkommenssituation und der Arbeitslosenquote weitere erklärende Faktoren wie den regionalen Besatz an Humankapital, die regionale Wettbewerbsfähigkeit, die Siedlungsstruktur und interregionale Pendlerverflechtungen als alternativen Angleichungsmechanismus für Arbeitsmarktungleichgewichte. Neben einer aggregierten Migrationsgleichung werden auch disaggregierte Modelle für verschiedene Altersgruppen geschätzt. Die Modelle werden als dynamische Paneldatenverfahren unter Verwendung von Instrumentenvariablen-Verfahren geschätzt, wobei ein besonderes Gewicht auf der Überprüfung der Schätzgüte der Modelle liegt. Neben Tests auf Endogenität der Regressoren hinsichtlich des Fehlerterms des Modells wird so auch für Abhängigkeiten über Querschnittseinheiten hinweg kontrolliert.

Sowohl die aggregierten als auch die altersgruppenspezifischen Resultate bestätigen den durch das neoklassische Migrationsmodell vorhergesagten Zusammenhang zwischen interner Migration und regionalen Arbeitsmarktungleichgewichten. Dabei gehen sowohl regionale Einkommensunterschiede als auch Disparitäten in den Arbeitslosenquoten als statistisch signifikant und dem Vorzeichen nach entsprechend der erwarteten Wirkungsrichtung des neoklassischen Modells in die Schätzgleichung ein. Die Resultate für die Arbeitsmarktvariablen bleiben auch dann stabil, wenn weitere Erklärungsgrößen wie der regionale Humankapital-Besatz, die regionale Wettbewerbsfähigkeit und Unterschiede in der Siedlungsstruktur in das empirische Modell aufgenommen werden. Die Hinzunahme von interregionalen Pendlerverflechtungen in die Migrationsgleichung zeigt eine negative Korrelation zwischen den beiden Variablen, was erste Rückschlüsse auf die substitutive Beziehung zwischen diesen Ausgleichsmechanismen für Arbeitsmarktungleichgewichte zulässt. Eine gesteigerte regionale Wettbewerbsfähigkeit hat den erwarteten positiven Einfluss auf Netto-In-Migrationsströme. Die empirischen Ergebnisse verdeutlichen auch die Bedeutung von regionalen Heterogenitäten im Migrationsverhalten, gemessen anhand von Unterschieden in der Siedlungsstruktur und den beiden West-Ost-Makroregionen (diese werden einerseits approximiert durch die Verwendung von fixen Effekten für einzelne Bundesländer sowie durch einen aggregierten Dummy für Ostdeutschland andererseits). Allerdings zeigen die Ergebnisse auch, dass der Effekt der zentralen Arbeitsmarktvariablen durch die Hinzunahme weiterer Variablen in das Modell nicht beeinflusst wird. In Analogie zu früheren empirischen Ergebnissen bestätigt die Untersuchung somit, dass makroregionale Unterschiede zwar von Bedeutung sind, diese jedoch nicht zu einer systematischen Ablehnung des neoklassischen Migrationsmodells führen. Die disaggregierten Ergebnisse für verschiedene Altersgruppen zeigen, dass insbesondere junge Alterskohorten zwischen 18 und 25 sowie 25 und 30 Jahren stark auf Unterschiede im regionalen Lohnniveau und der Arbeitslosenquote reagieren.

Um diesen Zusammenhang näher zu untersuchen, werden abschließend statistische Maße zur Bestimmung der relativen Bedeutung der Arbeitsmarktvariablen in einem multiplen Regressionsmodell unter Verwendung weiterer Erklärungsgrößen berechnet. Die Resultate zeigen, dass für junge Kohorten bis zu 65% der Variation in den innerdeutschen Migrationsströmen durch die beiden Arbeitsmarktgrößen Einkommens- und Arbeitslosenquotenunterschiede erklärt werden können. Diese Ergebnisse unterstreichen insgesamt die prominente Rolle von regionalen Arbeitsmarktkonditionen bei der Bestimmung von Migrationsbewegungen der Erwerbsbevölkerung in Deutschland.

Acknowledgements We thank Björn Alecke for helpful comments regarding the proper estimation of regional migration functions and help with the dataset.

Appendix

Baseline and augmented regression results by age groups

Table 8 Baseline migration model based on system GMM estimation

$nm_{ij,t}$	To18	18to25	25to30	30to50	50to65	Over65
$nm_{ij,t-1}$	0.87***	0.86***	0.86***	0.87***	0.90***	0.88***
	(0.001)	(0.005)	(0.004)	(0.002)	(0.001)	(0.002)
$\tilde{u}_{ij,t-1}$	−0.78***	−0.91***	−0.78***	−0.42***	0.19***	−0.03
	(0.044)	(0.156)	(0.148)	(0.036)	(0.019)	(0.018)
$\tilde{y}_{ij,t-1}$	0.28**	3.73***	4.03***	0.25**	−0.83***	−0.59***
	(0.112)	(0.406)	(0.395)	(0.102)	(0.042)	(0.043)
BW	−0.31***	−0.35***	−0.37***	−0.17***	0.11***	0.01
	(0.035)	(0.093)	(0.093)	(0.018)	(0.016)	(0.011)
BAY	−0.28***	−0.21***	−0.20***	−0.15***	0.07***	−0.01
	(0.031)	(0.075)	(0.077)	(0.018)	(0.016)	(0.009)
BER	0.42***	1.67**	1.32	0.12	−0.17***	−0.02
	(0.144)	(0.721)	(0.937)	(0.187)	(0.054)	(0.068)
BRA	0.59***	0.89***	1.12***	0.36***	−0.24***	−0.06***
	(0.044)	(0.171)	(0.156)	(0.052)	(0.019)	(0.018)
BRE	−0.06	1.95***	−0.38	−0.03	0.04	−0.10***
	(0.256)	(0.610)	(0.470)	(0.161)	(0.107)	(0.133)
HH	−0.11	−0.12	−1.22	−0.12	0.07	0.09
	(0.410)	(0.712)	(1.133)	(0.018)	(0.125)	(0.160)
HES	−0.18***	−0.22*	−0.27**	−0.12***	0.09***	0.03
	(0.045))	(0.133)	(0.110)	(0.018)	(0.031)	(0.027)
MV	0.48***	1.11***	1.19***	0.26***	−0.31***	−0.12***
	(0.047)	(0.171)	(0.164)	(0.051)	(0.022)	(0.021)
NIE	−0.01	0.14**	0.15**	−0.02	−0.05***	−0.04***
	(0.020)	(0.065)	(0.057)	(0.017)	(0.011)	(0.007)
NRW	−0.01	0.08	0.13*	−0.02	−0.01	−0.01
	(0.035)	(0.065)	(0.071)	(0.019)	(0.010)	(0.008)
RHP	−0.14***	0.15	0.08	−0.08***	0.02	−0.04***
	(0.035)	(0.102)	(0.089)	(0.017)	(0.026)	(0.014)
SAAR	0.46	0.49	2.20**	0.07	0.11	0.03
	(0.384)	(0.764)	(1.062)	(0.153)	(0.176)	(0.082)
SACH	0.47***	1.33***	1.49***	0.24***	−0.33***	−0.15***
	(0.055)	(0.194)	(0.177)	(0.052)	(0.028)	(0.022)
ST	0.53***	1.06***	1.17***	0.25***	−0.35***	−0.15***
	(0.088)	(0.177)	(0.178)	(0.051)	(0.020)	(0.021)
SH	0.10***	0.18*	0.19***	0.07***	0.07***	0.03
	(0.030)	(0.094)	(0.056)	(0.013)	(0.013)	(0.007)
TH	0.39***	1.42***	1.31***	0.21***	−0.34***	−0.18***
	(0.058)	(0.212)	(0.173)	(0.048)	(0.019)	(0.018)
No. of obs.	1067	1067	1067	1067	1067	1067
Time dummies (11)	Yes	Yes	Yes	Yes	Yes	Yes

Note: ***, **, * = denote significance levels at the 1, 5 and 10% level, respectively. BW = Baden-Württemberg, BAY = Bavaria, BER = Berlin, BRA = Brandenburg, BRE = Bremen, HH = Hamburg, HES = Hessen, MV = Mecklenburg-Vorpommern, NIE = Lower Saxony, NRW = North Rhine-Westphalia, RHP = Rhineland-Palatine, SAAR = Saarland, SACH = Saxony, ST = Saxony-Anhalt, SH = Schleswig-Holstein, TH = Thuringia.

Table 9 Augmented migration model based on system GMM estimation

$nm_{ij,t}$	To18	18to25	25to30	30to50	50to65	Over65
$nm_{ij,t-1}$	0.86***	0.85***	0.87***	0.87***	0.90***	0.84***
	(0.002)	(0.006)	(0.006)	(0.003)	(0.002)	(0.003)
$\tilde{u}_{ij,t-1}$	−1.10***	−0.72***	−0.84***	−0.72***	−0.02	−0.27***
	(0.117)	(0.239)	(0.256)	(0.061)	(0.032)	(0.035)
$\tilde{y}_{ij,t-1}$	−0.23	3.13***	5.28***	1.55***	−1.12***	−0.53***
	(0.175)	(0.633)	(0.369)	(0.157)	(0.097)	(0.090)
COMM	−0.10***	−0.06***	−0.04**	−0.01**	−0.02***	−0.03***
	(0.010)	(0.014)	(0.015)	(0.005)	(0.002)	(0.003)
BW	−0.19	−0.28	−0.85***	−0.39***	0.14***	−0.02
	(0.136)	(0.229)	(0.179)	(0.068)	(0.046)	(0.037)
BAY	−0.59***	−0.37	−0.98***	−0.39***	0.05	−0.11**
	(0.193)	(0.261)	(0.237)	(0.077)	(0.056)	(0.052)
BER	1.41***	1.02	0.81	0.59**	0.02	0.49***
	(0.481)	(1.182)	(1.157)	(0.279)	(0.136)	(0.186)
BRA	0.59***	0.37	065*	0.71***	−0.18***	0.04
	(0.164)	(0.365)	(0.350)	(0.103)	(0.046)	(0.055)
BRE	1.95**	2.76	−1.37	0.24	0.08	0.39
	(0.782)	(2.015)	(0.934)	(0.458)	(0.211)	(0.435)
HH	1.00	1.07	−1.23*	−0.41	0.35	0.09
	(1.173)	(1.183)	(0.629)	(0.424)	(0.368)	(0.611)
HES	−0.18	−0.33	−0.86***	−0.39***	0.13**	0.01
	(0.209)	(0.248)	(0.198)	(0.072)	(0.058)	(0.057)
MV	0.26*	0.41	0.76**	0.63***	−0.16***	−0.02
	(0.133)	(0.288)	(0.312)	(0.084)	(0.048)	(0.059)
NIE	−0.26*	−0.17	−0.52**	−0.06	0.05	−0.08**
	(0.139)	(0.264)	(0.198)	(0.083)	(0.047)	(0.033)
NRW	0.06	0.09	−0.12	−0.05	0.03	0.01
	(0.076)	(0.183)	(0.157)	(0.056)	(0.032)	(0.028)
RHP	−1.31***	−0.71***	−0.91***	−0.32***	−0.09*	−0.38***
	(0.226)	(0.247)	(0.286)	(0.089)	(0.051)	(0.066)
SAAR	−0.11	0.17	0.86	−0.33	0.26	0.06
	(0.736)	(1.279)	(1.361)	(0.488)	(0.249)	(0.227)
SACH	0.57***	0.96**	1.21***	0.75***	−0.34***	−0.08
	(0.188)	(0.405)	(0.403)	(0.115)	(0.061)	(0.066)
ST	−0.23	0.13	0.54	0.56***	−0.31***	−0.23***
	(0.176)	(0.321)	(0.352)	(0.088)	(0.048)	(0.055)
SH	0.11	−0.22	−0.56***	−0.02	0.09**	0.06
	(0.165)	(0.266)	(0.211)	(0.089)	(0.046)	(0.043)
TH	−0.45*	0.46	0.77**	0.53***	−0.34***	−0.18*
	(0.256)	(0.306)	(0.360)	(0.102)	(0.067)	(0.102)
No. of obs.	873	873	873	873	873	873
Time dummies (11)	Yes	Yes	Yes	Yes	Yes	Yes
Settlement type (6)	Yes	Yes	Yes	Yes	Yes	Yes

Note: ***, **, * = denote significance levels at the 1, 5 and 10% level, respectively. BW = Baden-Württemberg, BAY = Bavaria, BER = Berlin, BRA = Brandenburg, BRE = Bremen, HH = Hamburg, HES = Hessen, MV = Mecklenburg-Vorpommern, NIE = Lower Saxony, NRW = North Rhine-Westphalia, RHP = Rhineland-Palatine, SAAR = Saarland, SACH = Saxony, ST = Saxony-Anhalt, SH = Schleswig-Holstein, TH = Thuringia.

References

Alecke, B., Untiedt, G.: Determinanten der Binnenwanderung in Deutschland seit der Wiedervereinigung. Working Paper. University of Münster (2000)

Alecke, B., Untiedt, G.: Pendlerpotential in den Grenzregionen an der EU-Außengrenze. Methoden, Ergebnisse und Prognosen. In: Riedel, J., Untiedt, G. (Eds.) EU-Osterweiterung und deutsche Grenzregionen. ifo dresden studien 28/II. Ifo Institut für Wirtschaftsforschung e.V., München (2001)

Alecke, B., Mitze, T., Untiedt, G.: Internal migration, regional labour market dynamics and implications for German East–West-disparities – results from a panel VAR. Jahrb. Regionalwiss./Rev. Reg. Res. **30**(2), 159–189 (2010)

Arellano, M., Bond, S.: Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Rev. Econ. Stud. **58**, 277–297 (1991)

Baltagi, B.: Econometric Analysis of Panel Data. 4th edn. Wiley, Chichester (2008)

Baltagi, B., Bresson, G., Pirotte, A.: Panel unit root tests and spatial dependence. J. Appl. Econom. **22**(2), 339–360 (2007)

BBSR: Laufende Raumbeobachtungen des Bundesinstitutes für Bau-, Stadt- und Raumforschung. Various issues. Bundesamt für Bauwesen und Raumordnung, Bonn (2009)

Blundell, R., Bond, S.: Initial conditions and moment restrictions in dynamic panel data models. J. Econometrics **87**, 115–143 (1998)

Büchel, F., Schwarze, J.: Die Migration von Ost- nach Westdeutschland – Absicht und Realisierung. Mitt. Arbeitsmarkt. Berufsforsch. **27**(1), 43–52 (1994)

Brücker, H., Trübswetter, P.: Do the best go west? An analysis of the self-selection of employed East–West Migrants in Germany. IZA Discussion Papers 986. Forschungsinstitut zur Zukunft der Arbeit (IZA), Bonn (2004)

Burda, M.C.: The determinants of East–West German migration. Some first results. Eur. Econ. Rev. **37**, 452–461 (1993)

Burda, M.C., Hunt, J.: From reunification to economic integration: productivity and the labour market in East Germany. Brookings Pap. Eco. Ac. **2**, 1–92 (2001)

Burda, M.C., Härdle, W., Müller, M., Werwartz, A.: Semiparametric analysis of German East–West migration intentions: facts and theory. J. Appl. Econom. **13**, 525–541 (1998)

Daveri, F., Faini, R.: Where do migrants go? Oxford Econ. Pap. **51**(4), 595–622 (1998)

Decressin, J.: Internal migration in West Germany and implications for East–West salary convergence. Weltwirtsch. Arch. **130**, 231–257 (1994)

Destatis: Zu- und Fortzüge nach Altersgruppen (über die Kreisgrenze) – regionale Tiefe: Kreise und kreisfreie Städte. Wanderungsstatistik. Various issues. German Statistical Office, Wiesbaden (2009)

Devillanova, C., Garcia-Fontes, W.: Migration across Spanish provinces: evidence from the social security records (1978–1992). Investig. Econ. **28**(3), pp.461–487 (2004)

Dreger, C., Erber, G., Glocker, D.: Regional measures of human capital in the European Union. IZA Discussion Paper No. 3919. Forschungsinstitut zur Zukunft der Arbeit (IZA), Bonn (2009)

Eichenbaum, M., Hansen, L., Singelton, K.: A time series analysis of representative agent models of consumption and leisure under uncertainty. Q. J. Econ. **103**, 51–78 (1988)

Etzo, I.: Determinants of interregional migration in Italy: a panel data analysis, MPRA Paper No. 5307. University Library of Munich, Munich (2007)

Everaert, G., Pozzi, L.: Bootstrap-based bias correction for dynamic panels. J. Econ. Dyn. Control **31**(4), 1160–1184 (2007)

Evers, G.: Simultaneous models for migration and commuting: macro and micro economic approaches. In: van Dijk, J. et al. (Eds.) Migration and Labour Market Adjustment. Springer, Dordrecht (1989)

Fachin, S.: Long-run trends in internal migrations in Italy: a study on panel cointegration with dependent units. J. Appl. Econom. **51**(4), 401–428 (2007)

Federal Employment Agency: Arbeitsmarktstatistik der Bundesanstalt für Arbeit. Kreisdaten. Various issues. Available at: http://statistik.arbeitsagentur.de (2009). Accessed: July 2010

Ghatak, S., Mulhern, A., Watson, J.: Inter-regional migration in transition economies: the case of Poland. Rev. Dev. Econ. **12**(1), 209–222 (2008)

Greenwood, M.J., Hunt, G.L., Rickman, D.S., Treyz, G.I.: Migration, regional equilibrium, and the estimation of compensating differentials. Am. Econ. Rev. **81**, 1382–1390 (1991)

Grömping, U.: Relative importance for linear regression in R: the package relaimpo. J. Stat. Softw. **17**(1), 1–27 (2006)

Hansen, L.: Large sample properties of generalised method of moments estimators. Econometrica **50**, 1029–1054 (1982)

Harris, J.R., Todaro, M.P.: Migration, unemployment and development: a two sector analysis. Am. Econ. Rev. **60**, 126–142 (1970)

Hatzius, J.: Regional migration, unemployment and vacancies: evidence from West German microdata. Applied Economics Discussion Paper Series Nr.**164**, University Oxford (1994)

Hunt, J.: Staunching emigration from East Germany: age and the determinants of migration. J. Eur. Econ. Assoc. **4**(5), 1014–1037 (2006)

Im, K., Pesaran, M., Shin, Y.: Testing for unit roots in heterogeneous panels. J. Econometrics **115**, 53–74 (2003)

Jackman, R., Savouri, S.: Regional migration in Britain: an analysis of gross flows using NHS Central Register data. Econ. J. **102**(415), 1433–1450 (1992)

Kiviet, J.: On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. J. Econometrics **68**(1), 53–78 (1995)

Levin, A., Lin, C., Chu, C.: Unit root tests in panel data: asymptotic and finite-sample properties. J. Econometrics **115**, 53–74 (2002)

Maza, A., Villaverde, J.: Interregional migration in Spain: a semiparametric approach. Rev. Reg. Stud. **34**(2), 156–171 (2004)

Napolitano, O., Bonasia, M.: Determinants of different internal migration trends: the Italian experience. MPRA Paper No. 21734. University Library of Munich, Munich (2010)

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

Parikh, A., Van Leuvensteijn, M.: Interregional labour mobility, inequality and wage convergence. Appl. Econ. **35**, 931–941 (2003)

Pesaran, M.H.: Estimation and inference in large heterogeneous panels with a multifactor error structure. Econometrica **74**(4), 967–1012 (2006)

Pesaran, M.H.: A simple panel unit root test in the presence of cross-section dependence. J. Appl. Econom. **22**(2), 265–312 (2007)

Pissarides, C., McMaster, I.: Regional migration, wages and unemployment: empirical evidence and implications for policy. Oxford Econ. Pap. **42**, 812–831 (1990)

Puhani, P.A.: Labour mobility – an adjustment mechanism in Euroland? Ger. Econ. Rev. **2**(2), 127–140 (2001)

Rainer, H., Siedler, T.: Social networks in determining migration and labour market outcomes: evidence from the German reunification. Econ. Transit. **17**(4), 739–767 (2009)

Roodman, D.: How to Do xtabond2: An introduction to 'difference' and 'system' GMM in Stata. Stata J. **9**(1), 86–136 (2009)

Sarafidis, V., Wansbeek, T.: Cross-sectional dependence in panel data analysis. MPRA Paper No. 20367. University Library of Munich, Munich (2010)

Sarafidis, V., Robertson, D.: On the impact of error cross-sectional dependence in short dynamic panel estimation. Economet. J. **12**(1), 62–81 (2009)

Sarafidis, V., Yamagata, T., Robertson, D.: A test of cross section dependence for a linear dynamic panel model with regressors. J. Econometrics **148**(2), 149–161 (2009)

Sargan, J.: The estimation of economic relationships using instrumental variables. Econometrica **26**, 393–415 (1958)

Schwarze, J.: Beeinflusst das Lohngefälle zwischen Ost- und Westdeutschland das Migrationsverhalten der Ostdeutschen? Allg. Stat. Arch. **80**(1), 50–68 (1996)

Schwarze, J., Wagner, G.: Abwanderung von Arbeitskräften und Einkommenspolitik. DIW Wochenbericht **59**, 58–61 (1992)

Sevestre, P., Trognon, A.: Dynamic linear models. In: Matyas, L., Sevestre, P. (Eds.) The Econometrics of Panel Data. A Handbook of the Theory with Applications. 2nd edn. Springer, Dordrecht et al. (1995)

Silva, S., Hadri, K., Tremayne, A.: Panel unit root tests in the presence of cross-sectional dependence: finite sample performance and an application. Economet. J. **12**(2), 340–366 (2009)

Sjaastad, L.: The costs and returns of human migration. J. Polit. Econ. **70**, 80–93 (1962)

Suedekum, J.: Selective migration, union wage setting and unemployment disparities in West Germany. Int. Econ. J. **18**(1), 33–48 (2004)
Swiaczny, F., Graze, P., Schlömer, C.: Spatial impacts of demographic change in Germany. Urban population processes reconsidered. Z. Bevölkerungswiss. **33**, 181–205 (2008)
Todaro, M.: A model for labor migration and urban unemployment in less developed countries. Am. Econ. Rev. **59**(1), 138–148 (1969)
VGRdL: Volkswirtschaftliche Gesamtrechnungen der Bundesländer (Regional Accounts for German States). Available at: https://vgrdl.de (2009). Accessed: July 2010
Wagner, G.: Arbeitslosigkeit, Abwanderung und Pendeln von Arbeitskräften der Neuen Bundesländer. Soz. Fortschr. **41**(4), 84–89 (1992)
Westerlund, O.: Employment opportunities, wages and international migration in Sweden 1970–1989. J. Regional Sci. **37**, 55–73 (1997)

Timo Mitze is currently a researcher at the *Rheinisch-Westfälisches Institut für Wirtschaftsforschungxte* (RWI) in Essen. He studied economics in Münster and Stockholm. After his graduation he first joined the *Gesellschaft für Finanz- und Regionalanalysen* (GEFRA) in Münster, where he worked in the field of regional policy evaluation at national and EU levels. Subsequently, he received a PhD grant from the *Evangelisches Studienwerk Villigst* e.V. and took up a position at RWI in 2009. His main fields of research activity cover topics in empirical regional science, new economic geography, labor market and innovation economics with a special focus on panel and spatial econometrics.
e-mail: Mitze@rwi-essen.de

Janina Reinkowski is currently a researcher at the *ifo Institut für Wirtschaftsforschung* in Munich. She studied economics in Passau and Tours. After graduating, she first joined the *Rheinisch-Westfälisches Institut für Wirtschaftsforschung* (RWI) in Essen, where she mainly worked in the field of innovation policy analysis. Recently, she moved to the Department of Social Policy and Labor Markets at *ifo*. Her research interests cover applied labor market topics with a special focus on microeconometric program evaluation.
e-mail: Reinkowski@ifo.de