Original Paper

Tanja Fendel*

Migration and Regional Wage Disparities in Germany

DOI 10.1515/jbnst-2015-1001
Received April 23, 2014; revised October 9, 2014; accepted January 28, 2015

Abstract: This study provides an analysis of the effect of migration and commuting on regional wage disparities in Germany. Using the INKAR dataset and the GSOEP from the years 1998 to 2009, dynamic GMM panel estimations are applied to consider dynamics as well as the simultaneity between migration and regional labor market circumstances. To begin with, the influence of migration on relative wage levels is analyzed. The results show a small positive wage effect due to the overall regional migration balance. However, only domestic migration is relevant for analyzing the influence of migration on regional wage disparities. The wage effect due to domestic migration turns out to be smaller and negative. Regions seem to benefit primary from a combination of internal and foreign migration, however effects are small. Assuming that individuals usually move to high-wage regions, the negative wage effect of German migration would trigger an adjustment mechanism of wage disparities. Therefore, a second dynamic GMM panel estimation tests whether an influence of the regional wage levels on migration exists. Results show no statistically significant effects. An adjustment of existing wage disparities due to migration is not likely to occur in Germany in the next few years.

Keywords: migration, regional disparities, regional labor market, Germany
JEL classification: C23, C61, E24, R23

1 Introduction

The magnitude of labor market disparities among German regions is almost as large as the one between Germany and other countries. There are significant differences between East and West Germany resulting from the economic reconstruction in the East after the fall of the Berlin Wall. Furthermore, labor market circumstances vary on a regional level. Between 1998 and 2009, the average wage

***Corresponding author: Tanja Fendel,** Institut für Arbeitsmarkt und Berufsforschung – IAB, Regensburgerstr, 104, D- 90478 Nürnberg, E-mail: Tanja.Fendel@iab.de

level per hour ranged from about EUR 7 in regions of Mecklenburg-Western Pomerania to EUR 15 and above in some regions of Baden-Württemberg. During this time, the disparity between certain regional wages and the national mean rose for many regions. In some regions of Schleswig-Holstein, Thuringia, Bremen, and North Rhine-Westphalia below-average wages declined significantly. In contrast, there are regions, predominantly in Hesse, Bavaria and Baden-Württemberg, in which above-average wages increased. Average unemployment rates range from below 5 % in regions of Bavaria or Baden-Württemberg to above 20 % in some regions of Mecklenburg-Western Pomerania and Saxony. However, they remained relatively constant during this time.

There is no direct connection between regional labor market differences and the welfare level of the population in a particular region – especially since varying regional living costs reduce nominal wage differences. Many authors emphasize that economic disparities are inefficient. Taylor (1996) points out that they reduce national output and raise inflationary pressure. Elhorst (2003) adds that reducing disparities produces substantial social benefits. An important branch of research focuses on inter-regional migration as a mechanism for influencing labor market disparities; however, empirical evidence is mixed. It is questioned whether the influence is economically significant. Furthermore, the direction of the effect is unclear.

Traditional migration approaches assume that labor mobility reduces regional labor market disparities and can function as an adjustment mechanism. Möller (2001) describes different aspects of a regional adjustment process. When a region is hit by a severe adverse shock production is depressed and unemployment rises. High unemployment rates undermine the bargaining power of unions and individual employees. The wage pressure decreases and wages fall. Firms are able to offer lower prices. Consumption and – in the long run – labor demand increases. Furthermore, high emigration rates to other, more attractive regions will reinforce this process. Relative labor supply and unemployment decline.

Niebuhr et al. (2012) challenge traditional theories with new economic approaches such as New Economic Geography (NEG) established by Krugman (1991) and others. Increasing disparities may occur due to selective migration. The role of self-selection within the migration decision is an often emphasized issue (see Borjas 1987). Chiswick (1978) explains positive self-selection with the migration benefit which is higher for more able and higher motivated individuals. Hence, it is usually assumed that migrants are on average more educated or better skilled than individuals who choose to remain in their place of origin (Long 1973; Greenwood 1975; Chiswick 2000; Dustmann/Glitz 2011). In addition, economic migrants move to regions with above-average wage levels and low unemployment rates. Since additional high incomes are then spent in these regions, migration is assumed to lead to further prosperity. Furthermore, economically depressed regions with a net loss of the (highly skilled) population will suffer from a decreased demand for locally produced goods and services. Hence, disparities may increase due to migration.

This paper analyzes the relationship between migration and regional wage disparities in Germany for the period 1998 to 2009. It is examined whether migration and commuting contribute to increasing or decreasing disparities or have no significant economic influence. The analysis is based on the assumption of simultaneity between the regional labor market outcome and respective migration rates. It can be assumed that migration does not only influence the local labor market conditions but *is* also influenced by them. Attractive conditions may be an incentive for workers to move.

In the first step, a wage equation is estimated focusing on the influence of migration and commuting on the relative wage level of the "German Spatial Planning Regions" (Raumordnungsregionen, ROR). Dynamic GMM panel estimations are conducted to account for various simultaneities, for dynamic wage adjustments and for endogeneity due to regional fixed effects using data from the "Indicators and Maps on Urban Development in Germany and Europe" (Indikatoren und Karten zur Raum- und Stadtentwicklung in Deutschland und in Europa, INKAR) and the "German Socio-Economic Panel" (GSOEP). An error correction model provides a reconciliation of short-run and long-run effects of mobility on relative wages. The results show a significant small positive effect of regional migration on relative wage levels for both time frames. When the migration balance increases by 10 percentage point, the relative wage level increases by 0.0107 % in the long run. German regions seem to benefit from new citizens; however the effect is small. The migration balance considers both, internal migration within Germany and migration between Germany and other countries. When differentiating between Germans and transnational migration and only considering domestic migration, the effect on wages turns out to be smaller and negative. Hence, the above estimated positive effect only results from the combination of foreign and domestic migration.

The negative effect of internal migration would trigger an adjustment process when assuming that domestic migrants usually move to high-wage regions (although the impact is low due to the small estimated effect). Disparities lead to migration which decreases disparities. It is conceivable that employment prospects are important determinants of migration in Germany. Previous migration studies have argued that differences in economic opportunities between the source and destination regions due to income and unemployment rates are the main forces of migration. However, individual, family- and housing-specific factors highly influence the migration probability as well (Windzio 2004).

In the second step, a migration equation is estimated to analyze the influence of regional labor market circumstances on the domestic migration balance. Results indicate no effect of relative wage levels on migration. An adjustment process during which existing wage disparities decrease due to migration is not likely to occur in Germany in the next few years. However, the estimated positive effect of overall migration (both transnational and domestic) gives information about wage developments due to migration on a regional basis. In the last part of this study, German regions are named in which an already existing high gap between regional wages and the national mean is expected to increase in the coming years.

The outline of the paper is as follows: The second section reviews the relevant literature. In Section 3 the identification strategy is described. Section 4 presents the empirical analysis. It describes the econometric specification and presents the data set and the estimation results. Section 5 concludes.

2 Related literature

A lot of research has been carried out on labor mobility and inter-regional disparities. A classification can be reached by envisioning the issue's simultaneity. The existing literature is divided into studies focusing on migration as the dependent variable and approaches that focus on the labor market outcome as the dependent variable.

Many previous studies take the traditional assumption "migration reduces disparities" as given. Here, it is questioned whether the effect of relative labor market circumstances on migration is significant enough to function as an adjustment mechanism. Decressin (1994) estimates the influence of shocks and regional differences of unemployment rates and salaries on migration flows in West Germany during the 1980s. Using a Least Squares Dummy Variables (LSDV) model with dummies for each region, he estimates a migration elasticity of 1.3 resulting from local salary increases (relative to those paid in another region) and of 3 due to increases of the local unemployment rate. His results show that the increase of unemployment in all regions contributes to an economically significant decrease of gross migration. In times of recession, migration is less likely to work as an adjustment mechanism. In contrast, Pissarides and McMaster (1990) show in their study on Great Britain between the years of 1961 and 1982 that this adjustment process worked so slowly that usually "compensating differentials" pre-exist. Applying a LSDV model in a first step, estimated effects of the regional wage level (unemployment rate) in relation to the average wage level (unemployment rate) on migration flows turn out to be low. With an adjustment equation, a second step estimation finds significant positive effects of relative unemployment rates on relative wage levels.

In light of Mundell's "Theory of Optimum Currency Areas" Puhani (2001) assumes that a high degree of factor mobility will be conducive to the success of Europe's currency union. Using Ordinary Least Squares (OLS) estimation, he estimates a migration elasticity with respect to the number of unemployed people in West German regions of only 0.00809 for the years 1985 to 1996. With respect to the regional GDP at purchasing power parities (which is used instead of wages due to data availability), he estimates a migration elasticity of only 0.00136. Hence, from a ceteris paribus 1 % change in the GDP of a West German region results a migration-induced population increase of 0.00136 %. He concludes that the degree of labor mobility in major European nation states appears to be too low to act as an adjustment mechanism at least in the short run. Focusing on West Germany, Windzio (2004) analyzes main moving determinants of migration flows between the superior more attractive south and the inferior north region. By using a three-level model, he considers individual and regional factors as well as the respective time frame. Following his results, the moving probability is lower for individuals living in regions with high unemployment rates and higher for individuals with an academic degree. Arntz (2005) analyzes main work-related moving motives of Germans with regard to qualifications for the years 1975 to 2001. Using a two- level nested logit model, she considers individual as well as regional destination-specific factors. Her results indicate that the skill composition of job flows highly matters for the relevance of respective determinants. While highly skilled job movers are responsive to regional wage differences, unemployment differences only influence the migration decision of less skilled workers. Furthermore, migration costs appear to decrease with education: the proportion of highly skilled movers increases significantly with migration distance.

Due to the historical background, many examinations on German migration focus on mobility from the Eastern to the Western part after the reunification. Considering West-East migration, net migration from East to West Germany between 1989 and 2007 amounted to 1.7 million people (Wolff 2009). Since 1990 data on individuals from Eastern Germany is included in the GSOEP. Considering the first two waves, Burda (1993) estimates binomial logit models with the dependent variable taking the value 1 when a person living in the East can imagine to move to West Germany or to West Berlin. While wages and wage increases appeared not to have any effect in these early years, a person's age has a negative influence on the migration desire. Using a switching regression model and data from the IAB-employment sample, Brücker and Trübswetter (2004) find that migrants are positively selected with respect to unobserved abilities. Due to the longer observation period compared to Burda (1993), their results indicate that wage differentials and differences in employment opportunities were main forces of East-West migration after the German reunification.

Using the GSOEP waves 1990, 1991, 1996 and 1997, Hunt (2000) analyzes the determinants of emigration or commuting from East to West Germany with multinomial logit models. Her results indicate that migrants are on average younger and more skilled than non-migrants. Furthermore, using data on level of German Bundesländer from the years 1991 to 1996, she estimates the effect of regional labor market conditions on East to West migration with OLS regression. She explains a downward trend in East to West moving relative to within-west migration with the observed wage convergence during this time period. Hunt (2006) confirms these results. Again using data on level of German Bundesländer from the years 1991 to 2000, she estimates a fixed effects model and shows that rising wages reduce Eastern emigration. Furthermore, the SOEP indicates that young people are more sensitive to wage differences while older people are more sensitive to unemployment rates. Her results also show self-selection within the migration decision.

Regarding simultaneity, some studies focus on explaining the labor market situation with migration. Molho (1995) points out the dynamics of an adjustment process. His results show that for the UK in 1981, higher unemployment rates in remote areas can be explained by low out-migration rates caused by *distance deterrence* (and *cumulative inertia*). This is, however, only characterized by the sum of the distance from region i to other regions. Molho (1995) assumes that people in inaccessible areas who have fewer out-migration opportunities, stay unemployed for longer periods. Østbye and Westerlund (2007) estimate a neo-classical growth model with a System GMM estimator to consider endogeneity of the migration variable. The effect of migration on changes in GDP per capita is examined for Norwegian and Swedish counties over the time period 1980–2000. They show that migration adds to convergence of local labor market outcomes in Sweden but reduces convergence in Norway. However, convergence is measured by comparing the coefficient of the lagged dependent variable when it is a) controlled for migration and b) not controlled for migration. Therefore, the study does not measure convergence between regions, but over time.

Comparable to this study, Niebuhr et al. (2012) analyze the effects of labor mobility on German labor market disparities. A dynamic panel GMM estimation is conducted to consider simultaneity between mobility and the regional labor market outcome. The authors suggest that mobility decreases disparities, but only in unemployment rates. Wage disparities appear to remain unaffected. Interpretations are based on very strong assumptions. For a wage equation instead of wages in relation to the national mean, absolute wage rates are used. Although there is no per se connection to the relative wage position, an estimated negative migration effect on the absolute wage level is interpreted as a negative migration effect on regional wage disparities. This implies that migration mainly takes place from low- to high- wage regions. For the unemployment equation the relative unemployment rate is used but conclusions are also based on strong assumptions. A negative influence of the migration rate on the relative unemployment rate is understood as a negative effect on unemployment disparities. Only when people significantly react on regional labor market disparities by moving, the provided interpretation should be true. Otherwise results only give information about migration effects on the regional labor market outcome. Therefore, after estimating a relative wage equation, in a second step this study estimates a migration equation. Comparable to the previous studies in this section, it is analyzed whether migration rates of German regions are significantly influenced by the respective labor market outcome. In this case, migration may contribute to a process which may influence regional labor market disparities in the long run.

3 Identification strategy

In the first step of the empirical analysis the effect of migration on regional nominal wage levels is estimated. The regional value in relation to the national mean is used to consider the effect on the relative regional wage position. The relative regional wage level $\left(\frac{w_{it}}{w_t}\right)$ equals the average wage level of region i w_{it} divided by the average national wage level w_t in time t. The same applies for the considered relative regional unemployment rate $\left(\frac{u_{it}}{u_t}\right)$. In the following model, no distinction will be drawn between migration and commuting. The considerations are based on balance rates. The migration balance is the difference between a region's in- and out-mobility rates as the share of the regional population. Here, the estimation is carried out for the overall migration balance and the migration balance which only considers internal migration, ignoring migration between Germany and other countries. In the second step, a migration equation is estimated to analyze the influence of labor market disparities on domestic migration.

A simultaneous equation model (SEM) as given in eqs [1] and [2] describes theoretically the simultaneity between the wage and migration equation. However, in the empirical analysis the two equations are estimated separately, considering simultaneity by instrumenting possible endogenous variables as described below.

$$\left(\frac{w_{it}}{w_t}\right) = f\left(migB_{it}; \left(\frac{u_{it}}{u_t}\right); R_{it}; X'_{it}; \alpha_i; \tau_t^w; \varepsilon_{it}\right) \quad [1]$$

$$migB_{it} = g\left(\left(\frac{w_{it}}{w_t}\right); \left(\frac{u_{it}}{u_t}\right); R_{it}; \eta_i; \upsilon_t; e_{it}\right) \quad [2]$$

The wage eq. [1] describes the relation between different regional factors and the relative regional wage level of region i in t $\left(\frac{w_{it}}{w_t}\right)$. With regard to the research aim of this study, the effect of migration $migB_{it}$is central here and the estimation aims at determining whether the effect is positive or negative. Furthermore, the relative regional unemployment rate $\left(\frac{u_{it}}{u_t}\right)$ and the regional average rent per square meter R_{it} are assumed to be main determinants. While the relative unemployment rate is expected to have a negative influence, the average rent is assumed to have a positive influence on the local nominal wage level. Influencing regional varying living costs, local rents are usually relevant for salary negotiations.

Displayed by the vector X'_{it}, four additional regional control variables are integrated in eq. [1]. First, the share of qualified employees is considered. The qualification structure of the local labor force is of main importance to the relative average wage level of a region. Second, a region's population density per square kilometres is assumed to influence the wage level. It is mainly a proxy for a region's degree of urbanization. The migration effect on wages probably varies between urban and rural areas and especially in regions with a high share of highly skilled workers. Therefore, it is important to control for these factors. Third, the share of women and fourth of self employed individuals is considered in X'_{it} due to an assumed strong influence on the local relative wage level. All remaining factors influencing the relative wage level of a region and for a certain time will be represented by regional effects α_i, time effects τ_t^w and by a structural error term ε_{it}.

The causal interpretation of the migration eq. [2] is related to the decision to move to region i in time t. Pissarides and McMaster (1990) derive a migration function on the basis of the migration probability of a single household. A household or a person moves when the gross gain from moving exceeds their costs. Costs depend on observable and unobservable individual characteristics which are randomly distributed among the population. Therefore, the migration balance is a positive function of the gross gain from moving to a region. Following John Hicks (1932), local differences in net economic advantages are the main determinants of moving gains. In the model the relative wage level $\left(\frac{w_{it}}{w_t}\right)$, the relative unemployment rate $\left(\frac{u_{it}}{u_t}\right)$ and the regional average rent per square meter R_{it} are assumed to provide economic incentives to move to a region. A positive relation between wages and migration can be expected. A high unemployment rate is expected to reduce the employment opportunities of migrants and would therefore create a deterring effect. The same applies for the regional average rent per square meter. Various empirical studies estimate the effect of regional push and pull factors on migration. In traditional appraisals, the regional unemployment rate and the income or average nominal wage level are usually considered as main determinants (see Ritchey 1976; Greenwood 1975 for an overview). However, Renas and Kumar (1978) argue that nominal money income variables lead to a misspecification of the migration equation when local cost of living variables are not included. They show that variables measuring costs of living and the rate of change of these costs significantly influence migration. Empirical studies exist that measure the relevance of housing costs as a main component of the regional living costs. Cebula (2002) estimates a significant negative influence of the housing price index on the population change rate of U.S.-states. Furthermore, Pack (1973) shows that the lack of appropriate accommodation reduces the attraction of a city for migrants.

On the personal level, there are other main determinants of the moving gain such as year of birth, marital status and educational level. These determinants do not vary much over time and are treated as regional effects represented by η_i (Pissarides/McMaster 1990). In addition, time effects, displayed by υ_t, can be assumed due to varying migration patterns over time. All remaining factors influencing the migration rate will be represented by a structural error term e_{it}.

When assuming simultaneity between eqs [1] and [2], the observed data does not represent the amount of migrant people exogenously influencing the endogenous wage variable. Neither does the data represent a given exogenous wage level influencing the endogenous migration rate exogenously. It is conceivable that regional migration rates do not only influence labor market outcomes but are also influenced by them. Individuals may move to regions with attractive labor market circumstances and thus influence the respective labor market circumstances. Furthermore, simultaneity can be assumed for the rent variable in both equations. Regarding eq. [1], the local nominal wage may not only determine housing costs. Usually regional housing costs also influence wage negotiations. Regarding eq. [2], rents do not only work as incentive or deterrent factor within the migration decision. Due to the influence on housing demand, migration rates also influence regional housing costs.

It can be shown that the migration variable in eq. [1] is generally correlated with the structural error of eq. [1] if $\left(\frac{w_{it}}{w_t}\right)$ has a significant influence on $migB_{it}$ in eq. [2] and that the wage variable in eq. [2] is correlated with the structural error of eq. [2] if $migB_{it}$ has a significant influence on $\left(\frac{w_{it}}{w_t}\right)$ in eq. [1]. Therefore, if simultaneity exists, the migration variable is endogenous in eq. [1] and the wage variable is endogenous in eq. [2] and the rent and unemployment variable are endogenous in both equations. In this case, an OLS estimation suffers from simultaneity bias (Wooldridge 2003). To identify the equations and to solve the problem of endogenous explanatory variables, instrumental variables are needed for eq. [1] as well as for eq. [2]. The estimation method is introduced in Section 4.1.

Estimating the wage equation, either a positive or negative effect of migration on the relative wage level will result. The respective effect can be explained within the neoclassical labor market model. Here, a distinction has to be drawn between labor supply and labor demand effects. The labor supply in a region increases due to a positive migration balance. Yet, when considering selective migration, the in-flow of qualified workers may raise productivity and may increase labor demand as well (Niebuhr 2012). Unfortunately the given data does not indicate the qualification composition of the migration flows. However, there is a broad agreement in literature that migration is selective. Hence, it can be assumed that both, the labor demand and supply curve will shift to the right. The extent of these shifts will determine the wage effect. As presented in Figure 2(a) (see electronic appendix at www.jbnst.de/en) a negative effect of migration on wages will result due to a higher shift of the supply curve. A positive effect will result due to a distinct right shift of the labor demand curve (see Figure 2(b) in the electronic appendix at www.jbnst.de/en).

In the long run, labor demand may increase due to more consumption of new high-income receivers. Consumption and resulting wage increases due to commuting should be lower than due to migration. It can be assumed that commuters spend more money at their places of residence rather than in their work region (Elhorst 2003).

The central research question of this study is whether migration increases regional wage disparities. Since only disparities between German regions and not between Germany and other countries are observed, the effect of internal migration on wages is central for the research question. How does a positive or negative wage effect of domestic migration relates to disparities? The second estimation analyzes whether or not a high relative wage level and/or a low unemployment rate provide incentives to move to a certain region. In the literature it is often assumed that workers vote with their feet and move to high-wage/low-unemployment regions (Borjas 2000). If internal migration is significantly triggered by the local wage level, a positive migration effect on wages will lead to increasing disparities resulting in a self-reinforcing process: High disparities lead to migration and further increase disparities. A negative migration effect on wages will lead to a reduction of disparities. Disparities, which lead to migration, are likely to result in a wage adjustment process. When people or households do not move due to regional wage differences, migration is not likely to have a significant effect on disparities.

4 Empirical analysis

4.1 Econometric specification

The same framework is used to estimate the influences of regional migration rates on relative wage levels and to estimate the effects of regional labor market factors on migration rates. In the next passage, there will be a detailed description of the model for estimating migration effects followed by a brief specification on the second issue.

To estimate percentage changes, the dependent variable in the wage equation is given in logarithm. Following Harvey (1981), a dynamic model is carried out which differentiates between a short-run and a long-term effect. An integrated error correction mechanism ensures that the variables follow a steady-state growth path.

In the third section, the wage equation with main determinants is introduced in eq. [1]. An implicit linear approximation of this function is given by:

$$ln\left(\frac{w_{it}}{w_t}\right)^{*} = \beta_0 + \beta_2 migB_{it} + \beta_3 \frac{u_{it}}{u_t} + \beta_4 R_{it} + \beta_5 X'_{it} + \alpha_i + \tau_t \quad [3]$$

where $ln\left(\frac{w_{it}}{w_t}\right)^{*}$ is the nominal relative wage level of the structural equation. Depending on the specification $migB_{it}$ represents the overall migration balance, the domestic migration balance (MigB, DomMigB) and/or the commuting balance (ComB) in region i in time t. Main determinants of the local wage level are the relative regional unemployment rate $\frac{u_{it}}{u_t}$ and the average rent per square meter. The vector X'_{it} contains additional regional factors influencing the wage level. As justified in Section 3, X'_{it} includes the region's population density per square kilometres and the share of highly qualified employees, of women and of self-employed employees in a specific region. Time invariant regional effects are represented by α_i. Time effects which do not vary between regions are represented by τ_t.

The structural eq. [3] gives the long-run effect of migration on wages β_2. As Engle and Granger (1987) propose, economic series must be differenced before the assumption of stationarity holds. After taking first differences, the variables become so called co-integrated as represented in the adjustment eq. [4]:

$$\Delta ln\left(\frac{w_{it}}{w_t}\right) = (1-\beta_1)\left[ln\left(\frac{w_{it}}{w_t}\right)^{*} - ln\left(\frac{w_{i,t-1}}{w_{t-1}}\right)\right] + \varepsilon_{it} \qquad [4]$$

ε_{it} represents the structural disturbance term. Entering eqs [3] in [4] with $\Delta ln\left(\frac{w_{it}}{w_t}\right) = ln\left(\frac{w_{it}}{w_t}\right) - ln\left(\frac{w_{i,t-1}}{w_{t-1}}\right)$ and solving for $ln\left(\frac{w_{it}}{w_t}\right)$ gives the equation to be estimated:

$$ln\left(\frac{w_{it}}{w_t}\right) = \beta_0(1-\beta_1) + \beta_1 ln\left(\frac{w_{i,t-1}}{w_{t-1}}\right) + \beta_2(1-\beta_1)migB_{it} + \beta_3(1-\beta_1)\frac{u_{it}}{u_t} \qquad [5]$$
$$+ \beta_4(1-\beta_1)R_{it} + \beta_5(1-\beta_1)X'_{it} + (1-\beta_1)\alpha_i + (1-\beta_1)\tau_t + \varepsilon_{it}$$

where $\beta_2(1-\beta_1)$ yields the short-run effect of a relative wage level change in response to a 1% change of the migration rate. A series of abnormally large random disturbances influencing the wage development, may lead to a difference of the short-run and long-term effect of migration. After estimating the coefficient of the lagged dependent variable β_1 in eqs [5], [4] with the two terms multiplied by $(1-\beta_1)$ and ε_{it} drives the relative wage level back towards its long-run growth path. Therefore, resulting from the structural equation in eq. [3], $\beta_2 = [\beta_2(1-\beta_1)]/(1-\beta_1)$ represents the respective long-run wage effect of migration on wages. When carrying out an instrumental variable estimation, it has to be considered whether a correlation between $ln\left(\frac{w_{i,t-1}}{w_{t-1}}\right)$ and ε_{it} in first differences exists, leading to MA(1) errors.

This procedure is also used for estimating the long-run effect of labor market parameters on migration. The regional domestic migration balance as dependent variable is not given in logarithms because the variable takes on negative values. Hence effects can not be interpreted as elasticities. In the third section the migration equation with main determinants are introduced in eq. [2]. An implicit linear approximation of this function is given by:

$$migB_{it}^{*} = b_0 + b_2 ln\left(\frac{w_{it}}{w_t}\right) + b_3\left(\frac{u_{it}}{u_t}\right) + b_4 R_{it} + \eta_i + \upsilon_t \qquad [6]$$

where $migB_{it}^{*}$ is the domestic migration balance of the structural equation. η_i represents the regional effects and υ_t the time effects. A migration adjustment equation is given by:

$$\Delta migB_{it} = (1-b_1)\left[migB_{it}^{*} - migB_{i,t-1}\right] + e_{it} \qquad [7]$$

where e_{it} represents the disturbance term. Entering eqs [6] in [7] and solving for $migB_{it}$ gives the equation to be estimated:

$$migB_{it} = b_0(1-b_1) + b_1 migB_{i,t-1} + b_2(1-b_1)ln\left(\frac{w_{it}}{w_t}\right) + b_3(1-b_1)\left(\frac{u_{it}}{u_t}\right) \qquad [8]$$
$$+ (1-b_1)b_4 R_{it} b_4 + (1-b_1)\eta_i + (1-b_1)\upsilon_t + e_{it}$$

where $b_2(1-b_1)$represents the short-run effect of a change of the regional domestic migration balance in response to a one percentage point change of the relative wage level. Resulting from the structural equation in eq. [6], b_2 gives the long-run effect of the relative wage level. $b_3(1-b_1)$ gives the short-run effect of a change of the regional domestic migration balance in response to a one percentage point change of the relative unemployment rate. b_3 represents the respective long-run effect for the relative unemployment rate. Again for the estimation it has to be tested whether MA(1) errors exist.

The following passages discuss econometric issues considered in the estimation of eqs [5] and [8]. Heckman (1981) emphasizes that regression analysis may show a spurious effect when heterogeneity is not properly taken into account. Results would appear to demonstrate state dependence that does not exist. In addition to heterogeneity, Geyer and Steiner (2007) point out that there might exist unobserved serial correlation in time-varying error components and initial conditions or that relevant pre-sample history may be not taken properly into account in the estimation. With the given panel data it is possible to identify true state dependence.

The estimation method takes into account the regional fixed effects which are represented in eq. [5] by α_i and in eq. [8] by η_i. Following Jochimsen and Nuscheler (2011) it is plausible to treat the expected regional effects as fixed. They argue that there is no room for random effects (RE) when all regions of a country are included in the estimation. The Hausman test (see Arellano and Bond 1991) rejects the random effects specification. It can be assumed that regional effects and the considered explanatory variables are correlated. To identify true state dependence, the considered fixed effects and the disturbance term have to be conditionally uncorrelated. Since for neighbouring regions homoskedastic errors cannot be assumed, robust standard errors are computed. Year dummy variables are integrated into both eqs [5] and [8] to eliminate the assumed unobserved time effects τ_t and υ_t. Finally, there must be no remaining autocorrelation in the disturbance term which is tested with the Arellano Bond (1991) test.

To account for dynamics as described above, in both estimation equations the lagged depended variable is integrated as explanatory variable. Here a positive correlation with the regional fixed effects has to be considered. OLS estimation will lead to an inconsistent coefficient of the lagged dependent variable, especially since $T = 12$ is relatively small. An upwards biased “dynamic panel bias” will occur (Nickell 1981). As Bond (2002) shows, the fixed effects (FE) estimator does not eliminate the dynamic panel bias and is likely to be downward biased. He summarizes that a consistent estimator of the lagged dependent variable should lie between the FE and the OLS estimates, or should at least not be significantly lower than the former or significantly higher than the latter. Roodman (2009a) further emphasizes that a credible estimate should be below 1.00 since otherwise an unstable dynamic would occur.

To remove the fixed effects, first differences are taken. The lagged dependent variable is still potentially endogenous. However, while within the FE transformation instrumenting the lagged dependent variable with own lags is not possible, it is possible for the first difference transformation. In this study, the lagged dependent variable is not the only variable under consideration. When simultaneity exists between the local labor market outcome and migration as described in Section 3, the migration and unemployment variables in eq. [5] and the wage and unemployment variable in eq. [8] will also be endogenous. Furthermore, it is assumed that rents are endogenous in both equations. Due to an expected simultaneity bias, these variables should also be instrumented (see Wooldridge 2003). Since no exogenous instruments appear to exist (e. g. a source of exogenous variation in the migration variables that does not directly influence the relative wage level is very hard to find), the instruments used are again "internal", based on lags of the instrumented variables.

The Generalised Method of Moments (GMM) developed by Hansen (1982) leads to an asymptotically efficient estimator in this context. A one-step GMM estimator is not efficient when assuming heteroscedasticity. Hence, a two-step estimator with the Windmeijer bias correction is used (see Windmeijer 2005).

This study uses a Difference (DIF) GMM estimator and a System (SYS) GMM estimator. Within the DIF GMM estimator proposed by Arrelano and Bond (1991), first differences are taken and potentially endogenous variables are instrumented with usable lags of their own levels. Blundell and Bond (1998) show that past levels may convey little information about future changes, resulting in poor performance if the dependent variable is close to a random walk. The proposed SYS GMM estimator additionally estimates level (LEV) equations where endogenous variables are instrumented with own lagged differences. In this study, results of the SYS GMM estimator are generally preferred to results of the DIF estimator.

The System GMM estimator has superior finite sample properties in terms of bias and root mean squared error when series are persistent. However, Bun and Windmeijer (2009) show that instruments may still be weak. Using a concentration parameter proposed by Rothenberg (1984), they use a covariance stationary panel data AR(1) model to compare the information content of instruments in the difference and levels equation. When the variance of the idiosyncratic shocks $\left(\sigma_v^2\right)$ is larger than the variance of the unobserved heterogeneity term $\left(\sigma_\eta^2\right)$, the LEV model performs better in terms of a smaller concentration parameter, of a smaller LEV and SYS 2SLS bias and of a better Wald test performance. However, superiority of the SYS GMM estimator relative to the DIF one is usually shown for samples in which the variance of the regional effects is high relative to the variance of the transitory shock, such as $\left(\sigma_\eta^2\right) = \left(\sigma_\upsilon^2\right)$ or even $\left(\sigma_\eta^2\right) > \left(\sigma_\upsilon^2\right)$. In the former case $\left(\sigma_\eta^2\right) = \left(\sigma_\upsilon^2\right)$, concentration parameters and the distortion of the Wald test turn out to be equal for both the DIF and the LEV models. Therefore, as proposed by the authors, the weakness of the instruments used in this study is tested separately. An underidentification test is conducted to find out whether the used instruments are correlated with the endogenous regressors. The p-value of the Kleibergen-Paap (2006) rk LM statistic gives information on whether the excluded instruments are "relevant" and therefore whether the equation is identified. Furthermore, instruments are tested to see in how far they are only weakly correlated with the endogenous regressors by reporting the Kleibergen-Paap rk Wald F statistic. The Kleibergen-Paap Wald statistic is the robust counterpart of the Cragg-Donald Wald statistic. The critical values are the Stock Yogo (2005) IV critical values for the Cragg-Donald i.i.d. case. In this separate estimation for weak instrument testing, an endogeneity test of endogenous regressors will be applied as well (see Baum 2010).

The instruments should not only be sufficiently correlated with the included endogenous variables; the second necessary criterion for an instrument to be valid is exogeneity. Instruments should not be correlated with the error term and hence with the dependent variable. The Hansen J statistic is reported to test whether the instruments used are jointly valid in this respect (Hansen 1982). The Hansen test also examines whether the idiosyncratic disturbances follow a moving average process of first or higher order (MA(.)). Additionally, two Difference-in-Hansen tests are used to check the validity of a subset of instruments. While the Difference-in-Hansen test eq. [1] checks the validity of the subset of instruments based on the levels equation (which is only relevant for the System GMM estimation), the Difference-in-Hansen test eq. [2] checks the validity based on the dependent variable.

Although the Difference and System GMM estimator gained high popularity in the last years, they are not without problems. In addition to the weak instrument problem, instrument proliferation is another, related severe issue in the application of the GMM estimator to dynamic panel data models. The problems arising from instrument proliferation are well documented in the literature. However as Roodman (2009b) or Bontempi and Mammi (2012) point out, in connection with the Difference and System GMM estimator the issue needs to receive much more attention in research. Roodman (2009b) describes two main problems when the number of moment conditions is too large relative to the sample size. Firstly applying to instrumental variable estimators in general, instruments can overfit endogenous variables, resulting in a small sample bias in the direction of OLS. Secondly only applying to the two-step GMM estimator, estimates of the optimal weighting matrix tend to be very imprecise due to its high dimensions. As a result, the standard errors of the estimator tend to be severely downward biased and both Hansen tests can be greatly vitiated (Verbeek 2012). Results appear valid, creating J statistics with high p-values of 1.000 or close to 1.000. However, the implausible high p-values can be expected to result from instrument proliferation weakening the test's ability to detect a possible violation. Since there is no formal test or rule of thumb of how many instruments are too many, Roodman (2009b) proposes testing GMM results for robustness to reductions in the instrument set. As Windmeijer (2005) reports, Monte Carlo experiments showed that reducing the total number of instruments from 28 to 13 decreased the average bias of the two-step GMM estimator by 40 %.

With $T = 12$ and four or five potentially endogenous variables for each model of this study, there is a large number of instruments available. In this study, for the basic estimations the lag length of instruments is restricted to T-8. To detect a potential violation of the Hansen tests and to decrease the possible bias of instrument proliferation, the number of instruments is further reduced for the main results in two variants. The first variant restricts the instrument set to a one-lag period. The second one collapses the full instrument set into a smaller one by combining instruments through addition. An instrument matrix is squeezed horizontally and combines formerly distinct columns. The estimator will not separate empirical moments $\sum_{t,l} y_{i,t-l}\Delta e_{it}$ for each l and t but will only minimize the magnitude of the moments $\sum_{t} y_{i,t-l}\Delta e_{it}$ for each l. While the instrument count in the full set for Difference and System GMM is typically quadratic in T, the collapsing makes it – as the first variant above – linear in T (Roodman 2009b).

4.2 Data and descriptive statistics

The examination about the effect of migration and commuting rates of German regions on regional wage disparities is based on data from the "Indicators and Maps on Urban Development in Germany" (INKAR, Federal Institute for Research on Building, Urban Affairs and Spatial Development Bonn)[1] and from the "German Socio Economic Panel" (GSOEP)[2] from the years 1998 to 2009. The examination is

1 A description of INKAR can be downloaded from www.bbsr.bund.de/BBSR/DE/Veroeffentlichungen/INKAR/inkar_node.html.

2 A description of the GSOEP can be downloaded from www.diw.de/soep.

carried out on the regional level of the German Spatial Planning Regions (Raumordnungsregionen, RORs). RORs are classifications between German administrative districts and counties with a total of 96 regions. The INKAR dataset is supplied by the German Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR). It gives information on spatial development in several regional levels in the form of data, maps and charts. The data set provides a wide range of official regional figures about employment, population and age structure, mobility, education, living standards, environment, health, public budget, traffic and local economy in Germany. The GSOEP is supplied by the German Institute of Economic Research (DIW Berlin). It is a representative panel survey of private households in Germany. Among many other fields, it contains detailed information on household income and work hours.

SOEP data is used for differentiated wage information because the INKAR data set provides only information on gross earnings, not taking work hours into account. Average net hourly wages of SOEP respondents are merged to the INKAR data set on the regional level of RORs. For each ROR and year the number of wage observations in the SOEP for the considered time varies from 9 to 424. On average, the calculated wages are based on 100 observations. However, for RORs with few wage observations, the wages may not be fully representative and results may be biased to a certain extent. For further research it would be worthwhile to test results using other representative Data such as the employment history statistic of the IAB. For the following descriptive statistic part, more representative average wages are calculated using SOEP expansion factors. To consider dependencies between regions and for wages of one person over the observed years, averages are calculated using OLS estimation with robust standard errors. When estimating causal effects within the main regression analysis of this study, the use of sample weighting does not improve the results. The various econometric issues considered in the estimation already correct for heteroscedasticity and endogenous sampling (see Solon et al. 2013).

Wages are used in nominal form. Although the German federal statistical office is working on it, currently there is no representative data on regional differentiated prices available in Germany. As described in Section 3, in both estimated equations it will be controlled for local average rents. The BBSR supplies average rents, based on newspapers and internet advertisements.[3] Unfortunately, no data is available for years before 2004. However, for now and in contrast to SOEP information, it is the most representative available data base that includes all German regions. Therefore, average rents from 2004–2006 are used for the years before 2004 (BBSR-Wohnungsmarktbeobachtungssystem, IDN ImmoDaten GmbH).

3 More information on the BBSR-rents are available on http://www.bbsr.bund.de/BBSR/DE/Veroeffentlichungen/BBSROnline/2010/ON012010.html.

Figures 3 and 4 (see electronic appendix at www.jbnst.de/en) give information about labor market disparities in Germany. For each ROR averages for the years 1998 to 2009 are displayed over German “Bundesländer” (NUTS 1 level). Figure 3 depicts averages of net hourly wages. In West Germany, shown by the first 10 columns, average wages range from about EUR 9 (Landshut) to EUR 15 (Ostwürttemberg). In the Eastern Bundesländer, shown by the last 6 columns, wages range from about EUR 7 (Mecklenburgische Seenplatte) to EUR 11 (Havelland-Fläming). Figure 4 depicts disparities in local unemployment rates. Unemployment rates in the East range from 12 % (Südthüringen) to 21 % (Mecklenburgische Seenplatte). In the West they range from 5 % (Oberland) to above 14 % (Emscher-Lippe).

While unemployment disparities do not vary much during the observed time period, the distance of regional wage rates to the national mean increased in many regions. Figure 1 relates average relative wage rates from 1998–2001 with average relative wage rates from 2006–2009 for West and East Germany. Most Eastern wages (see chart a)) remain below the average in the observed years, with a relative wage lower than 1. There are eight regions above the diagonal indicating further relative decreases. In the West (chart b)) there are many regions in the upper-right quadrant with a relative wage higher than 1 indicating that wages remained above average. Here, points under the diagonal indicate further relative increase. Ostwürttemberg is a high outlayer. Further decreases of under average wage rates are higher than in the East.

The data set enables the observation of migration and commuting balances. With regard to migration, in addition to the overall migration balance, the rates for internal migration between different German regions are given. While migration flows are related to 100 residents of the respective ROR, commuting flows are related to 100 employees from the respective working population in the region. The mobility rates remain similar in the observed time period. Figure 5 (in the electronic appendix at www.jbnst.de/en) depicts the yearly change of the migration balance. Overall migration balances range from −1.38 to 1.76 %. Domestic migration balances range from −16.52 to 1.79 %. Commuting balances range from −43.91 to 30.44 %.

Additional regional control variables are used to estimate the influence of migration on wages and the effect of regional labor market circumstances on migration (see Section 4.1). Table 1 depicts the means and standard errors of the variables used in the estimation based on 96 RORs.

Figure 6 (see electronic appendix at www.jbnst.de/en) displays a scatter diagram relating relative wage levels (without using expansion factors) to migration balances. Although a distinct connection is not evident at first sight, a regression line illustrates a positive relation. Accordingly, wages in regions with high positive migration balances appear to be high.

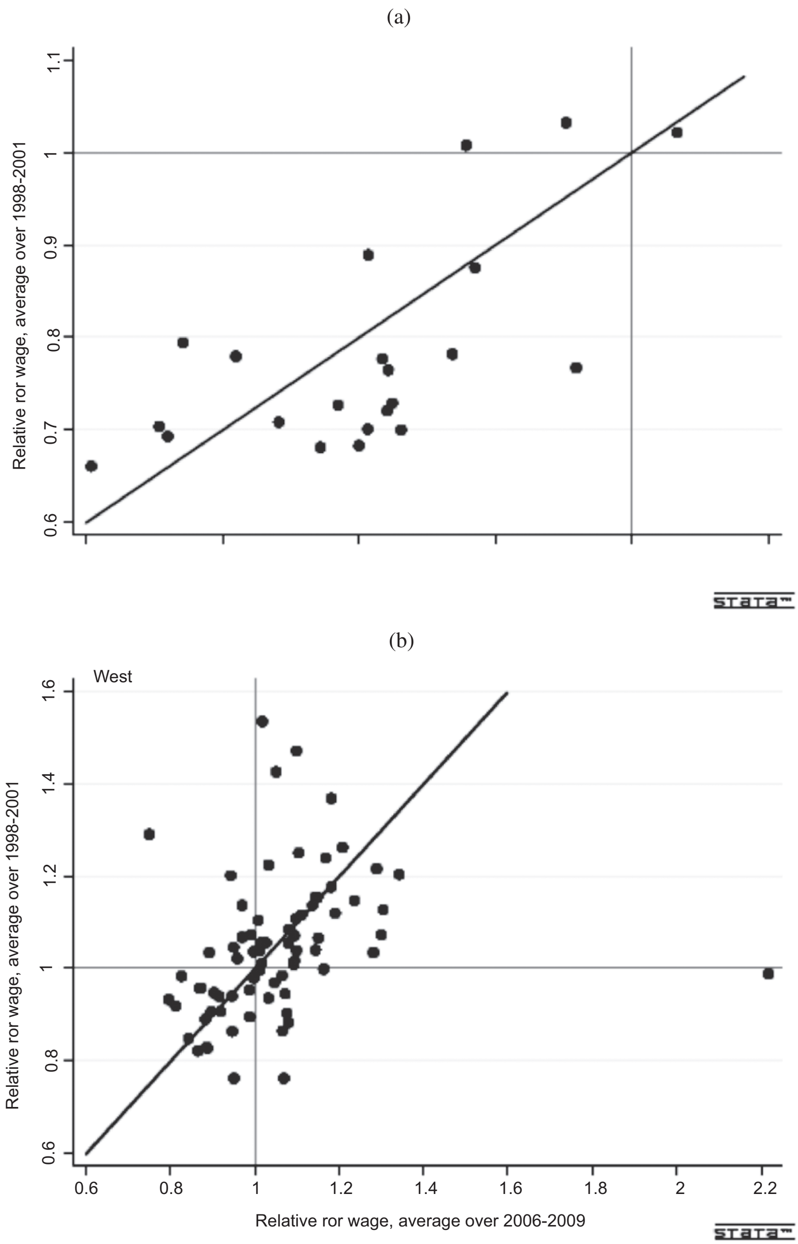

Figure 1: Change of relative wage levels from 1998–2001 to 2006–2009.
Source: SOEP.

Table 1: Mean and standard deviation of regional explanatory variables.

	Mean	Standard deviation
Wage per hour in EUR	10.23	(1.72)
Unemployment rate in %	10.04	(4.66)
Migration balance, related to residents	0.09	(0.43)
Within migration balance, related to residents	−0.06	(1.02)
Commuting balance, related to employees	−5.03	(11.67)
Rent per square meter in EUR	5.43	(0.85)
Population density, residents per km^2	330.57	(494.15)
Share of women in %, related to residents	50.99	(0.42)
Share of self-employed in %, related to employees	11.05	(1.95)
Share of highly qualified employees in %, related to employees	7.83	(2.80)

Source: INKAR, SOEP.

4.3 Results

In the following, results for the wage and migration equations are presented. In the basic estimations 4 lags are used (see Table 6, 7 and 8 in the electronic appendix at www.jbnst.de/en). Due to a large number of instruments resulting from four or five potentially endogenous variables, the *p*-values of the Hansen tests approach 1 in most of these outcomes. To reduce a potential bias of instrument proliferation, in the main results presented below, the instrument count is reduced in two variants as described at the end of Section 4.1. The instrument set is reduced to a one-lag period or into a smaller one by combining instruments through addition. For each model using four lags or one lag, a separate two-step feasible GMM estimation applies an underidentification and a weak instrument test as well as an endogeneity test of endogenous regressors.

The first part below presents results of the effect of migration on regional wage levels. Two specifications are used to estimate the effects of the migration and commuting balance. In the first specification, the migration rate considers both, migration between different German regions and between Germany and other countries. Analyzing the effect of migration on disparities, only internal migration should be taken into account. Therefore, in a second specification, the migration balance considers only migration between different German regions. The last part of this section presents results of the effect of regional wage levels and unemployment rates on the domestic migration balance.

4.3.1 Estimation of the wage equation considering over-all migration

Table 2 below summarizes the main results of the first specification. Results of the Difference GMM and System GMM equations are reported separately. The range of the FE and OLS point estimate of the lagged dependent variable is given by 0.3923–0.7377. While the coefficients of the Difference GMM estimation in Reg. 1a lies barely outside this interval, the remaining coefficients (Reg. 1b and 2a-c) lie in the interval. With 0.3306 the DIF GMM coefficient when using one lag is still very close to the given interval. However as discussed in Section 4.1, results of the SYS GMM estimator are in general preferred to results of the DIF GMM estimator when there are no contradicting indicators. While in Reg. 1a and 1b none of the mobility rates show significant coefficients, the migration balance in Reg. 2a and c show significant positive coefficients. Following results of the SYS GMM estimation of Reg. 2a, the effect of the overall migration balance on the relative wage level is small but positive and not only a short-term phenomenon. When the migration balance rate in a certain region increases by 10 percentage point (which equals a change in the value of 0.1), the relative wage level increases by 0.0042 % in the short run and by 0.0107 % in the long run (dividing 0.042 by [1–the coefficient of the lagged dependent variable]*0.1). As shown in Figure 5 (see electronic appendix at www.jbnst.de/en), yearly changes of the German migration balance in the observed years lie in an interval from −50 to +50 percentage points (which equals an interval of the value of −0.5 to +0.5). In these terms, the changes of the relative wage lie in an interval of −0.0535 to +0.0535 % per year. The commuting rate variable shows no significant effect in any of the outputs of this study.

All estimations of the first specification pass the autocorrelation test. To verify zero autocorrelation in first-differenced errors, serial correlation must not occur at any order higher than one. Using four lags (see Table 6 in the electronic appendix at www.jbnst.de/en) the Hansen and Difference-in-Hansen tests give high *p*-values of 1.000 which Roodman (2009b) calls the classical sign of instrument proliferation weakening the tests ability to detect a problem and to possibly biasing results. When reducing the instrument count, relevant coefficients stay significant and the values change only marginally. While results of Table 6 give approximate values, possibly suffering from overfitting bias, presented results in Table 2 are more reliable. Reducing the instrument count increases the ability of the Hansen test to detect a possible violation. Using only one lag (see Reg. 1a and 2a in Table 2) or using all lags collapsed (Reg. 1b and 2b in Table 2) lowers most *p*-values of the Hansen test to more realistic

Table 2: Estimation results for the wage equation –difference GMM and System GMM estimation with reduced lags, considering domestic and foreign migration (Specification 1).

	Reg. 1a	Reg. 1b	Reg. 2a	Reg. 2b	Reg. 2c
	(DIFF, 1 lag)	*(DIFF, all lags collapsed)*	*(SYSTEM, 1 lag)*	*(SYSTEM, all lags collapsed)*	*(SYSTEM, 1 lag, ComB excluded)*
ln (Rel. wage (t–1))	0.3306***	0.3980***	0.6071***	0.4397***	0.6091***
	(0.0573)	(0.0592)	(0.0523)	(0.0527)	(0.0508)
Rel. Unempl. Rate	0.0187	0.0069	−0.0293	−0.0066	−0.0403***
	(0.0638)	(0.0615)	(0.0278)	(0.0497)	(0.0177)
MigB	0.0257	0.0170	0.0421***	0.0229	0.0403**
	(0.0228)	(0.0269)	(0.0148)	(0.0217)	(0.0177)
ComB	−0.0033	−0.0021	0.0008	0.0029	
	(0.0046)	(0.0048)	(0.0009)	(0.0023)	
Additional Controlvar.	√	√	√	√	√
Year dummies	√	√	√	√	√
Obs	950	950	1045	1045	1045
Regions	95	95	95	95	95
Instruments	95	55	141	70	112
AR 1	0.000	0.000	0.000	0.000	0.000
AR 2	0.662	0.435	0.254	0.434	0.268
Hansen	0.288	0.592	0.999	0.248	0.658
Difference-Hansen (1)			1.000	0.224	0.897
Difference-Hansen (2)	0.147	0.664	1.000	0.338	0.992

Notes: * significant at 10 %, ** at 5 %, *** at 1 %. Robust standard errors are reported in parentheses.

Additional regional control variables: The average Rent, the population density, the share of women, of self-employed and of high-qualified employers.

Yearly time dummies are included in all regressions.

Reported estimates are based on the 2-step GMM estimator with the Windmeijer bias-correction.

The Arrelano-Bond test AR(1,2) is a test for first and second order autocorrelation in the first-differenced residuals.

The Hansen test is a test of the validity of overidentifying restrictions. The Difference-in-Hansen test checks the validity of the subset of instruments for the level equation (1) and of the subset based on the dependent variable (2). *P*-values are reported.

values. The same phenomenon applies for the Difference-in-Hansen tests. In Reg. 2a the *p*-values are still very high. Since five variables are instrumented, there are still 141 instruments in the System GMM estimation using only one lag. In order to further reduce the instrument count, the commuting variable is excluded in an additional regression (Reg. 2c). In this case the Hansen tests are passed at lower *p*-values. The used instruments appear to be valid.

For the separate GMM estimation of the first specification using one and four lags, the *p*-values of the Kleibergen-Paap rk LM statistic equal 0.000. This indicates that both equations are identified. Using four lags, the Kleibergen-Paap Wald rk F statistic for weak indentification testing equals 22.02 and 96.04 when using one lag. The F values being high enough, indicates that the used instruments are not weak. The endogeneity tests of endogenous regressors report *p*-values of 0.000 in both cases. Therefore, the hypotheses that respective variables are exogenous can be rejected.

4.3.2 Estimation of the wage equation considering only domestic migration

Table 3 summarizes the results for the second specification which only considers internal migration within Germany. The FE-OLS interval for the point estimate of the lagged dependent variable is given by 0.3913–0.7612. Again all estimates of the lagged relative wage level variable lie in or close around this interval. Although the respective coefficients of the Difference GMM estimation (0.3688 in Reg 3a and 0.3276 in Reg. 3b) lie outside this range, they are not significantly lower than the Fixed Effects estimate. The coefficients of the System GMM estimation lie with 0.6465 (Reg. 4a) and 0.4023 (Reg. 4b) inside the respective range. While in Reg. 4a the migration rate shows no statistically significant effect on wages, the respective coefficient of Reg. 4b and the coefficients of the Difference GMM estimation (Reg. 3a and b) are statistically significant. Hence in contrast to the first specification, the effect of internal migration on wages turns out to be zero, or very small and negative.

All regressions of the second specification pass the autocorrelation test. Using four lags in the output presented in Table 7 (see electronic appendix at www.jbnst.de/en), the *p*-values of the Hansen and Difference-in-Hansen tests reach 1.000 signalling instrument proliferation. Table 3 presents the respective results of the second specification when the instrument count is reduced. Coefficients equal their respective DIF and SYS estimation counterparts in value and significance. This gives evidence that a possible bias from instrument proliferation could be reduced. When reducing the instrument count for the Difference GMM estimation (Reg.3a and 3b), *p*-values of the Hansen and

Table 3: Estimation results for the wage equation – difference GMM and system GMM estimation with reduced lags, considering only domestic migration (Specification 2).

	Reg. 3a	Reg. 3b	Reg. 4a	Reg. 4b
	(DIFF, 1 lag)	*(DIFF, all lags collapsed)*	*(SYSTEM, 1 lag)*	*(SYSTEM, all lags collapsed)*
ln(Rel. wage (t–1))	0.3688***	0.3276***	0.6465***	0.4023***
	(0.0526)	(0.0652)	(0.0453)	(0.0558)
Rel. Unempl. Rate	−0.0191	−0.0312	−0.0470**	−0.0359
	(0.0492)	(0.0594)	(0.0247)	(0.0533)
DomMigB	−0.0066***	−0.0052**	−0.0002	−0.0061***
	(0.0016)	(0.0024)	(0.0020)	(0.0021)
ComB	−0.0002	0.0003	−0.0006	0.0038
	(0.0036)	(0.0041)	(0.0009)	(0.0030)
Additional Controlvar.	√	√	√	√
Year dummies	√	√	√	√
Obs	950	950	1045	1045
Regions	95	95	95	95
Instruments	95	55	141	70
AR 1	0.000	0.000	0.000	0.000
AR 2	0.495	0.624	0.190	0.465
Hansen	0.456	0.417	0.996	0.099
Difference-Hansen (1)			1.000	0.147
Difference-Hansen (2)	0.583	0.619	1.000	0.390

Notes: see Table 2.

Difference-in-Hansen eq. [2] tests decrease to more realistic values. Furthermore, the coefficient of the lagged dependent variable in Reg 3a increases in direction of the FE-OLS interval. When reducing the lag length to one lag in the SYS GMM estimation in Reg. 4a, the instrument count of 141 instruments is still very high and the *p*-values of the Hansen tests still equal or are close to one. When collapsing all lags, the instrument count is reduced to 70 instruments. However, in this case the main Hansen test is not passed. Some of the instruments do not appear to be fully exogenous in the SYS GMM estimation. This indicates that the Difference GMM estimation, proposing a small negative wage effect gives more accurate results for the second specification.

Results of the Difference GMM estimation using 1 lag (Reg. 3a) propose that when the domestic migration balance rate of a region increases by 10 percentage point (which equals a change in the value of 0.1), the relative wage level decreases by 0.00066 % in the short run and by 0.001 % in the long run. The positive effect of over-all migration on German regional wages seems to exist mostly due to mobility between Germany and other countries or due to the combination of domestic migration and transnational migration.

For this second specification using one and four lags, the reported Kleibergen-Paap rk LM statistic indicates that both models are identified. Furthermore, using four lags, the Kleibergen-Paap Wald rk F statistic for weak identification testing equals 22.02 and 96.04 when using one lag. Displayed by the F-values, instruments appear to be sufficiently correlated with the endogenous regressors. Results of the endogeneity tests of endogenous regressors indicate that respective variables are endogenous.

4.3.3 Estimation of the migration equation

A second-step estimation analyzes the influence of labor market disparities on migration within Germany. Here, the migration equation, given in eq. [8] is estimated using the domestic migration balance. Using DIF and SYS GMM estimations, it is accounted for dynamics of the migration rate and for endogeneity due to regional fixed effects. Furthermore, justified by the results of the former wage estimation, the model considers simultaneity. Table 4 depicts main results. Table 8 (in the electronic appendix at www.jbnst.de/en) depicts the results for the basic versions using 4 lags.

The credible FE-OLS range of the point estimate on the lagged dependent variable for this version is given by 0.8315–0.9092. The lagged dependent variable coefficients of the DIF GMM estimation (Table 4, Reg. 5a-b) and of the SYS GMM estimation (Reg. 6a-b) lie closely above the given FE-OLS interval; they are not significantly higher than the OLS coefficient. For the DIF GMM estimation, the unemployment and rent variables show significant negative effects. In both regressions the wage coefficient is not statistically significant. For the SYS GMM estimation, except for the coefficient of the lagged dependent variable, none of the remaining coefficients is statistically significant.

All regressions of the migration equation pass the autocorrelation test. Results of the basic estimations (see Table 8 in the electronic appendix at www.jbnst.de/en) show extremely high p-values for the Hansen and Hansen-in-Difference tests. Since the instrument count is high, instrument proliferation can be assumed to weaken the test's ability to evaluate instruments and to bias coefficients and standard errors. When reducing the instrument count in the regressions presented in Table 4, respective coefficients stay significant and their values change only marginally, possibly reducing an instrument proliferation bias. For the Difference GMM estimation (Reg. 5a-b) the Hansen test

Table 4: Estimation results for the migration equation, Difference GMM and System GMM estimation with reduced lags.

	Reg. 5a	Reg. 5b	Reg. 6a	Reg. 6b
	(DIFF, 1 lag)	*(DIFF, all lags collapsed)*	*(SYSTEM, 1 lag)*	*(SYSTEM, all lags collapsed)*
DomMig.B (t–1)	0.9301***	0.9787***	0.9180***	0.9607***
	(0.0595)	(0.0642)	(0.0103)	(0.0523)
ln(Rel. Wage)	−0.3570	−0.5609	0.1972	0.6813
	(0.4170)	(0.6803)	(0.2747)	(0.6753)
rel. Unempl.	−1.1470***	−1.1660***	−0.0312	−0.1424
	(0.4020)	(0.3322)	(0.0441)	(0.1347)
Rents	−0.4875***	−1.0037***	0.0154	−0.0761
	(0.2734)	(0.3484)	(0.0419)	(0.1001)
Year dummies	√	√	√	√
Obs.	950	950	1045	1045
Regions	95	95	95	95
Instruments	50	50	86	55
AR 1	0.007	0.002	0.001	0.007
AR 2	0.274	0.275	0.247	0.242
Hansen	0.016	0.023	0.117	0.026
Difference- Hansen (1)			0.912	0.634
Difference- Hansen (2)	0.332	0.495	0.752	0.951

Notes: see Table 2.

produces extremely low p-values indicating that some instruments may not be fully exogenous. Concerning the advantages of the SYS GMM estimator Verbeek (2012) emphasizes that if the true coefficient of the lagged dependent variable is close to unity which appears to be the case for the migration equation, lagged levels are poor instruments for first differences. When reducing the instrument count of the System GMM estimation (Reg. 6a and b), the Hansen and Difference-in-Hansen tests show lower p-values larger than 0.05 (except for the Hansen test in Reg. 6b). Again results of the SYS GMM estimator are preferred to results of the DIF GMM estimator, indicating that there is no effect of the relative wage level on the domestic migration balance.

For the separate GMM estimation using four lags, the underidentification test indicates problems. A p-value of the Kleibergen-Paap rk LM statistic larger than 0.05 (p-value = 0.0820) indicates that the model is underidentified. However with regard to the Kleibergen-Paap Wald rk F statistic of 152.73, instruments appear to be sufficiently correlated with the endogenous regressors.

The p-value of the Kleibergen-Paap rk LM statistic decreases to 0.012 when only using one lag. In this case the Kleibergen-Paap Wald rk F statistic equals 9.92; the value is still larger than the critical value of 5. Coefficients of the basic regressions using four lags may be biased due to a weak instrument problem. The problem is no longer apparent when the instrument count is reduced. Therefore, also in terms of weak instrument problems results of Reg. 5a, b and 6a, b appear to be more reliable. Results of the endogeneity tests of endogenous regressors indicate for both models with one or four lags that respective variables are endogenous.

Summarized as a main result of this section, the regional domestic migration balance does not appear to be influenced by the regional relative wage level. Households or individuals, who decide to relocate, seem to have other reasons for moving such as family or housing related issues. As Pissarides and McMaster (1990) point out, the regional fixed effects include these other determinants. Individual or household characteristics such as age, educational level or household composition influence the migration probability and may vary across regions.

5 Conclusion

This study provides an analysis of the effect of German regional mobility rates on regional wage disparities. The estimation uses data from the “Indicators and Maps on Urban Development in Germany and Europe” (INKAR) and the “German Socio Economic Panel” (GSOEP) from the years 1998 to 2009 on the basis of German Spatial Planning Regions (RORs). In a first step, a structural wage equation is carried out to estimate the effect of the regional migration and commuting balances on the relative regional wage level. Here, a first specification estimates the effect of the overall migration balance while a second specification considers only internal migration. The latter is relevant for analyzing the effect of migration on disparities between different German regions. All estimations of this study calculate short-run and long-term effects. Dynamic panel GMM estimations are conducted to account for simultaneity between the regional labor market situation and mobility, for dynamic wage adjustments and for endogeneity due to regional fixed effects.

Due to four or five endogenous variables in each estimation equation of this study, instrument proliferation can be assumed, also when the lag length is reduced to T-8 for basic estimations. Too many instruments may bias estimates and standard errors as well as may weaken the validity of the Hansen-Tests. The classical sign of too optimistic Hansen-test *p*-values are observed in the basic results. Following recommendations in the literature, for the main results of this study the instrument count is further reduced in two variants. In all cases this lowers the *p*-values to more realistic numbers and verifies basic results. However, the literature provides little guidance on how many instruments is too many. The research on instrument proliferation within the Difference and System GMM estimators is still in the early stages and as other authors claim, more methods are needed to further reduce the problem.

The results of the first specification indicate a small positive short-run and long-term wage effect of the regional overall migration balances. When the regional migration balance rate increases by 10 percentage point, the relative wage level increases by 0.0107 % in the long run. The yearly changes of the migration balance of German regions lie in an interval of −50 to +50 percentage points (which equals a change in the value of −0.5 to +0.5). Therefore, regions with a positive change rate of the migration balance can expect an increase of the relative wage rate of up to 0.054 %. Regions with a negative change rate of the migration balance are expected to deal with a respective negative effect. In sum, the regions seem to benefit from new citizens, although the effect is small. One reason for the positive effect may be a higher consumption demand. This might also provide an explanation for the fact that no wage effect of commuting could be found in any output of this study. Commuters may spend more money at their places of residence rather than in their work regions.

The results of the second specification show a small but negative wage effect of the migration balance which only considers migration between German regions. When the domestic migration balance in a region increases by 10 percentage points, the relative wage level decreases by 0.001 % in the long run. By comparing the results of specifications 1 and 2, it can be concluded that German regions seem to benefit mostly from the combination of domestic migration and foreign migration. However it has to be considered that both wage effects of overall and internal migration are very small.

In debates about migration, politicians often emphasize the fear that the regional labor market situation is worsened by migration, especially from other countries. The estimated small positive wage effect indicates that earning circumstances in German regions due not chance or may even improve with high positive migration balances. A possible explanation of the positive wage effect is selective migration. Self-selection can be considered in qualification but also in terms of unobserved characteristics such as general capability, motivation and courage. The moving incentive should be higher for individuals being more able or higher motivated in occupational field. This should especially apply with regard to the decision for relocating to another country. Hence, political debates about the impact of foreign migration should consider these results. As forecasted by the

BBSR, the German population will decrease by about 2.52 % in the years 2010 to 2030, the number of employed people is expected to decrease by about 6.69 %. According to the ifo Institute, Germany may compensate the increasing labor demand by the unemployed people in the short run. However, labor reserves will not be able to fulfill the labor demand of firms in the long run. While for Germany as a whole, the GDP is supposed to increase by 1.14 % p.a. from 2010 to 2030, growth rates will highly differ on the regional level. The ifo Institute emphasize the demographical development as a main reason for the differences. Regions with a stable labor force potential will be better off (Berlemann et al. 2012). This stability may also be supported by foreign migration.

Migration literature assumes that domestic migrants move to regions with high wages and low unemployment rates. Focusing on work-related moving, Arntz (2010) confirms this assumption for Germany from 1975 to 2001. She shows that high-skilled job movers are mainly motivated by regional wage differences while less-skilled job movers are influenced by unemployment differences. As shown in studies presented in the second section, wage differences also mainly influence the migration flows from East to West Germany. It can be assumed that these crossing-boarder moves are primarily economically motivated moves. For the purpose of the research question, this study questions in a second step whether economic factors are main determinants of *general* migration within Germany. Only in this case, the estimated wage effect due to migration would be able to influence and to change regional wage disparities in the long run. Results indicate that the relative wage level shows no statistically significant influence on the regional domestic migration balance. Results probably differ from previous studies since the estimation considers migration that is not only motivated by economic factors. On average German households or individuals who decide to move, seem to have other reasons for moving such as family or housing-related issues.

Traditional migration literature assumes that labor mobility reduces regional labor market disparities. This conclusion is also made by Niebuhr et al. (2012) concerning unemployment rates. In this study when only considering internal migration within German regions, the estimated migration effect on relative wage levels turns out to be small and negative. Indeed, a negative effect may lead to an adjustment of disparities (although the impact would be low due to the small effect), but only if the migration rate due to wage disparities is high enough. Without analyzing this relationship, results remain hypothetical. Results of the second main estimation of this study indicate that German migration is not influenced by regional relative wage levels. An adjustment mechanism where disparities trigger migration and by this reduce existing disparities is not likely to occur in the coming years. Regional bindings might deter individuals leaving German regions with unattractive labor market circumstances. In addition, regional living expenses vary across Germany and thereby reduce earning disparities (Goebel et al. 2009). In this sense, disparities may still be too low to provide moving incentives.

On a single case basis, results can be used to determine how much the relative wage levels of regions may change due to migration. Since the estimated effect of the regional over-all migration balance on wages is positive, wages in regions with a positive balance will increase and decrease in regions with a negative balance. Hence, disparities will increase when regions with above-average wage levels are confronted with increases of the migration balance. In addition, disparities increase when regions with under-average wage levels are faced with decreases of the migration balances. For the analysis relative average wage levels are calculated using SOEP expansion factors. Table 5 (see electronic appendix at www.jbnst.de/en) lists regions in which a positive wage distance to the national mean may further increase in the coming years. Large increases of above-average wage levels due to a high average growth of the migration balance can be expected for Hamburg, Neckar-Alb, Rhein-Neckar and Munich. The average change rate of the migration balance (for the years 1998 to 2009) in the first three regions is 2 percentage points which implies an average increase of the relative wage level of 0.002 % due to the estimated wage effect of 0.00107. The average change rate in Munich is 5 percentage points. Hence, an average increase of the relative wage level of 0.005 % can be expected.

Large decreases of under-average wage levels due to a large amount of relocations can be expected in particular for regions of Brandenburg. Here occur the highest negative average change rates of the regional migration balance. In Uckermark-Barnim, Prignitz-Oberhavel and Oderland-Spree average change rates (for the years 1998 to 2009) range from −11 to −14 percentage points. The estimated wage effect of 0.00107 leads to average declines of the relative wage level of 0.012 to 0.015 %.

Acknowledgements: I would like to thank Viktor Steiner, Beate Jochimsen, Davud Rostam-Afschar and Jan Goebel. Furthermore, I would like to thank Peter Winker and both referees for their important and insightful comments.

References

Arntz, M. (2010), What Attracts Human Capital? Understanding the Skill Composition of Interregional Job Matches in Germany. Regional Studies 44(4): 423–441.

Arntz, M. (2005), The Geographical Mobility of Unemployed Workers, Evidence from West Germany. ZEW Discussion Paper 05(34): 1–43.

Arrelano, M., S. Bond (1991), Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. Review of Economic Studies 58 (2): 277–297.

Baum, C.F., M.E. Schaffer, S. Stillman (2010), ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression. [stata help]. Erhältlich unter: http://www.ppge.ufrgs.br/epr/arquivos/help-ivreg2.pdf (Stand 2010, zuletzt aufgerufen am 17.09.2014).

Berlemann, M., J. Freese, M. Luik, J. Wesselhöft (2012), Prognose der wirtschaftlichen Entwicklung 2010 bis 2030. [Abschlussbericht zum Forschungsvorhaben „Verkehrssverflechtungsprognose 2030 sowie Netzumlegung auf die Verkehrsträger“. ifo Institut, Niederlassung Dresden]. Erhältlich unter: http://www.bmvi.de/SharedDocs/DE/Anlage/VerkehrUndMobilitaet/verkehrsverflechtungsprognose-2030-strukturdaten prognose-los-1.pdf?__blob=publicationFile (Stand 2012, zuletzt aufgerufen am 17.09.2014).

Blundell, R., S. Bond (1998), Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. Journal of Econometrics 87: 115–143.

Bond, S. (2002), Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice. Cemmap working paper CWP09/02, 1–36.

Borjas, G.J. (2000), Economics of Migration. International Encyclopedia of the Social and Behavioral Sciences SECTION 3.4(38): 1–22.

Borjas, G.J. (1987), Self-Selection and the Earnings of Immigrants. American Economic Review 77(4): 531–553.

Bontempi, M.E., I. Mammi (2012), A Strategy to Reduce the Count of Moment Conditions in Panel Data GMM. Munich Personal RePEc Arcive Paper 40720, 1–41.

Brücker, H., P. Trübswetter (2004), Do the Best Go West? An Analysis of the Self-Selection of Employed East-West Migrants in Germany. IZA Discussion Paper 986, 1–43.

Bun, M.J.G., F. Windmeijer (2009), The Weak Instrument Problem of the System GMM Estimator in Dynamic Panel Data Models. Econometrics Journal 13: 95–126.

Burda, M.C. (1993), The Determinants of East-West German Migration, Some First Results. European Economic Review 37: 452–461.

Cebula, R.J. (2002), Net Interstate Population Growth Rates and the Tiebout-Tullock Hypothesis: New Empirical Evidence, 1990–2000. Atlantic Economic Journal 30(4): 414–421.

Chiswick, B.R. (2000), Are Immigrants Favorably Self-Selected? An Economic Analysis. IZA Discussion Paper 131, 1–32.

Chiswick, B.R. (1978), The Effect of Americanization on the Earnings of Foreign-Born Men. Journal of Political Economy 86 (5): 897–922.

Decressin, J.W. (1994), Internal Migration in West Germany and Implications for East-West Salary Convergence. Weltwirtschaftliches Archiv 130(2): 231–257.

Dustmann, C., Glitz, A. (2011), Migration and Education. 327–429 in: E.A. Hanushek, S. Machin, L. Woessmann, (Hrsg.), Handbook of the Economics of Education 4. Amsterdam, North-Holland.

Elhorst, P. (2003), The Mystery of Regional Unemployment Differentials: Theoretical and Empirical Explanations. Journal of Economic Surveys 17(5): 709–748.

Engle, R.F., C.W.J. Granger (1987), Co-Integration and Error Correction: Representation, Estimation, and Testing. Econometrica 55(2): 251–276.

Geyer, J., V. Steiner (2007), Short-Run and Long-Term Effects of Childbirth on Mothers' Employment and Working Hours Across Institutional Regimes, An Empirical Analysis based on the European Community Household Panel. DIW Discussion Paper 682, 1–31.

Goebel, J., J.R. Frick, M.M. Grabka (2009), Preisunterschiede milder einkommensgefälle zwischen west und ost. Wochenbericht Des DIW Berlin 51–52: 888–895.

Greenwood, M.J. (1975), Research on Internal Migration in the United States: A Survey. Journal of Economic Literature 13(2): 397–433.

Hansen, L.P. (1982), Large Sample Properties of Generalised Method of Moments Estimators. Econometrica 50: 1029–1054.

Harvey, A. (1981), The Econometric Analysis of Time Series. Oxford, Philip Allan (second edition).

Heckman, J. (1981), Heterogeneity and State Dependence. 91–131 in: S. Rosen (Hrsg.), Studies in Labor Markets. Chicago, University of Chicago Press.

Hicks, J. (1932), The Theory of Wages. New York: Macmillan. St. Martins' Press.

Hunt, J. (2006), Staunching Emigration from East Germany: Age and the Determinants of Migration. Journal of the European Economic Association 4(5): 1014–1037.

Hunt, J. (2000), Why Do People Still Live in East Germany? NBER Working Paper Series 7564, 1–49.

Jochimsen, B., R. Nuscheler (2011), The Political Economy of the German Länder Deficits: Weak Governments Meet Strong Finance Ministers. Applied Economics 43(19): 2399–2415.

Kleibergen, F., R. Paap (2006), Generalized Reduced Rank Tests Using the Singular Value Decomposition. Journal of Econometrics 133: 97–126.

Krugman, P. (1991), Increasing Returns and Economic Geography. Journal of Political Economy 99(31): 483–499.

Long, L.H. (1973), Migration Differentials by Education and Occupation: Trends and Variations. Demography 10(2): 243–258.

Molho, I. (1995), Migrant Inertia, Accessibility and Local Unemployment. Economica, New Series 62(245): 123–132.

Möller, J. (2001), Regional Adjustment Dynamics. HWWA Discussion Paper 146, 1–44.

Nickell, S. (1981), Biases in Dynamic Models with Fixed Effects. Econometrica 49(6): 1417–1426.

Niebuhr, A., N. Granato, A. Haas, S. Hamann (2012), Does Labour Mobility Reduce Disparities Between Regional Labour Markets in Germany? Beiträge Zur Jahrestagung Des Vereins Für Socialpolitik F11-V2: 1–44.

Østbye, S., O. Westerlund (2007), Is Migration Important for Regional Convergence? Comparative Evidence for Norwegian and Swedish Counties, 1980–2000. Regional Studies 41(7): 901–915.

Pack, J.R. (1973), Determinants of Migration to Central Cities. Journal of Regional Science 13(2): 249–260.

Pissarides, C.A., I. McMaster (1990), Regional Migration, Wages and Unemployment: Empirical Evidence and Implications for Policy. Oxford Economic Papers, New Series 42(4): 812–831.

Puhani, P. (2001), Labour Mobility: An Adjustment Mechanism in Euroland? Empirical Evidence for Western Germany, France and Italy. German Economic Review 2(2): 127–140.

Renas, S.M., R. Kumar (1978), The Cost of Living, Labor Market Opportunities, and the Migration Decision: A Case of Misspecification? Annals of Regional Science 12: 95–104.
Ritchey, P.N. (1976), Explanations of Migration. Annual Review of Sociology 2: 363–404.
Roodman, D.M. (2009a), How To Do xtabond2: An Introduction to Difference and System GMM in Stata. Stata Journal 9(1): 86–136.
Roodman, D.M. (2009b), A Note on the Theme of Too Many Instruments. Oxford Bulletin of Economics and Statistics 71(1): 135–158.
Rothenberg, T.J. (1984), Approximating the Distributions of Econometric Estimators and Test Statistics. 881–935 in: Z. Griliches, M.D. Intriligator (Hrsg.), Handbook of Econometrics, Volume 2. Amsterdam, North-Holland.
Solon, G., S.J. Haider, J. Wooldridge (2013), What Are We Weighting For? NBER Working Paper 18859, 1–29.
Stock, J.H., M. Yogo (2005), Testing for Weak Instruments in Linear IV Regression. 80–108 in: D.W.K. Andrews; J.H. Stock, (Hrsg.), Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg. New York, Cambridge University Press.
Taylor, J. (1996), Regional Problems and Policies: A European Perspective. Australian Journal of Regional Studies 2: 103–131.
Verbeek, M. (2012), A Guide to Modern Econometrics. West Sussex, John Wiley & Sons Ltd (fourth edition).
Windmeijer, F. (2005), A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimator. Journal of Econometrics 126: 25–51.
Windzio, M. (2004), Zwischen nord- und süddeutschland: Die überwindung räumlicher distanzen bei der arbeitsmarktmobilität. Zeitschrift Für Arbeitsmarktforschung 37(1): 29–44.
Wolff, S. (2009), Determinants of East-West Migration in Germany: A Macroeconometric Analysis. Working paper, Georg-August-Universität Göttingen, 1–46.
Wooldridge, J.M. (2003), Introductory Econometrics, A Modern Approach. South-Western: Thomson.

Supplemental Material: The online version of this article (DOI: 10.1515/jbnst-2015-1001) offers supplementary material, available to authorized users.