CQE Center for Quantitative Economics

Regional labour migration – Stylized facts for Germany

Mark Trede † and Michael Zimmermann †

93/2020

† Department of Economics, University of Münster, Germany

wissen.leben
WWU Münster

Regional labour migration – Stylized facts for Germany*

Mark Trede† and Michael Zimmermann‡

University of Münster

December 2020

Abstract: We present stylized facts of local German labour markets using German administrative data and regional prices. We document regional variation in: unemployment rates, nominal wages, and real wages. The real wage gap between East and West Germany still persists 30 years after reunification whereas unemployment rates tend to converge. We investigate monthly worker flows across NUTS 3 regions. Unemployed workers in depressed regions are less likely to move than unemployed workers in prosperous regions. The most mobile group are unemployed workers in dense regions. Employed workers are less willing to move and have procyclical fluctuations in their moving rates.

JEL codes: R23, J61, J63, C55

Keywords: labour mobility; business cycle fluctuations; regional disparities

*Funding from DFG is gratefully acknowledged. This paper is part of the joint Franco-German research project *LOCLABMARKET* in collaboration with partners from Aix-Marseille School of Economics (AMSE), Zentrum für Europäische Wirtschaftsforschung (ZEW) and Centre national de la recherche scientifique (CNRS).

†Corresponding author. Institute for Econometrics and Economic Statistics, Department of Economics, University of Münster, Germany, Tel.: +492518325006, Email: mark.trede@uni-muenster.de

‡Institute for Econometrics and Economic Statistics, Department of Economics, University of Münster, Germany, Email: michael.zimmermann@wiwi.uni-muenster.de

1 Introduction

The finding that globalization does not only have positive effects everywhere but may indeed hurt some regions, has triggered a lot of research on local labour markets both in theoretical and empirical economics. Local shocks affect workers in different ways, for instance, by changes in the transition rates between labour markets status. Local labour market conditions are also related to the incentive to migrate between regions within the home state. While international migration regularly leads to heated debates, internal migration flows within states receive much less attention. International migration is more prominent in the public debate as it is sometimes claimed that it might have a negative impact on national security or labour markets, in particular for unskilled jobs. International migration is also more perceptible due to language barriers or cultural differences. But a careful investigation of internal migration flows also provides interesting insights. As in most other countries, internal migration is also an important phenomenon in Germany. The specific German history, especially the large and persistent regional economic division – the "Poor East" lagging behind the "Wealthy West" – makes it a particularly interesting object to study.

This paper's contribution is purely descriptive. We give a comprehensive account of different aspects of local labour markets in Germany between 1995 and 2017 from which we distill "stylized facts". The regional resolution is (roughly) at the level of "Landkreise" or NUTS 3. Our primary data source is the sample of integrated labour market biographies (SIAB) containing information about a 2% random sample for 328 regions. The data set allows insights into trends and patterns of regional migration and local transition rates between labour market status across Germany. Adding newly available regional price levels (Weinand and von Auer, 2019), we compute regional real wages and investigate their relationship with other local variables.

Stylized facts about regional worker migration are interesting for several reasons. First, they facilitate drawing a broad-brush picture of the historical development as excessive complexities are suppressed. They may also help put the economic situation into

a wider sociological or political context. Second, stylized facts enable us to assess the plausibility of economic theories without formal statistical testing. In general, economic theories are not capable to explain the observations in all details. Statistical hypothesis tests are thus bound to reject virtually any theoretical model if the number of observations is large. In contrast, by comparing model implications to the stylized facts one can judge if the model is useful for the problem at hand. For example, a plain economic theory would suggest that regions with high unemployment levels experience more net emigration than regions with fewer unemployed. However, this is in contrast to what the stylized facts reveal. Hence, there is need to develop other, more sophisticated economic theories, e.g. postulating persistent differences in unemployment levels even in equilibrium (as in e.g. Molho, 1995 or Aragon et al., 2003). Third, the stylized facts are relevant for policy makers. Government efforts to reduce regional unemployment differentials are pointless when the economy is already in equilibrium. If, however, the differentials are transitory and reflect an ongoing convergence process, there is scope for government intervention to speed up convergence.

Thirty years after the fall of the wall between East and West Germany, migration from the East to the West continues, causing fears that East Germany will remain poorer and less innovative. Glorius (2010) and Hunt (2000) point out that East Germany suffers from brain drain. While a certain degree of mobility of skilled workers is desirable, East Germany lacked immigration of a similar scope until recently. Up to now, East Germany has lost more than 2m residents, approximately 10% of the population of the former German Democratic Republic. The claim that unemployment differentials will level out over time is supported by the stylized fact that the gap between East and West unemployment rates is (slowly) narrowing. The related literature focuses mainly on aggregate disparities between East and West Germany. Smolny (2011) and Heiland (2004) are among the first contributions investigating state-level data (NUTS 1). They point out that there is not only a gap between East and West Germany but also between North and South Germany. They argue that interregional migration should be investigated more thoroughly. A recently published paper by Bauer et al. (2019)

analyses internal migration behaviour in Germany. They find out that labour market variables have powerful explanation for internal migration. In particular, migration flows of younger age cohorts are attracted to urban areas by these factors.

The remainder of the paper is structured as follows: The next section provides information about the SIAB dataset and the data source for the regional price levels. Section 3 presents aggregated stylized facts without a fine-grained regional disaggregation. In section 4, we derive regional stylized facts. A descriptive regression model of regional mobility is outlined in section 5. Section 6 concludes.

2 The Data

Our analysis primarily relies on the Sample of Integrated Labour Market Biographies (SIAB) provided by the Institute for Employment Research (IAB).[1] The SIAB is a 2% random sample of integrated labour market biographies in Germany from 1975 to 2017 which consists of all German residents who (i) have jobs that are subject to social security, (ii) are in marginal part-time employment, (iii) receive unemployment or social benefits, (iv) are registered as job seekers, or (v) participate in employment or training programs. Civil servants and self-employed workers are not included in the dataset.[2] In total, the dataset covers about 80% of the German labour force.

Each spell in the dataset provides daily information on individuals' employment status, region of work and average daily (pre-tax) wage or social security benefits, as well as the socio-economic variables gender, age, education and nationality. The SIAB's main data source is the employment history (Beschäftigten-Historik, BeH) collected by the Employment Agency for administrative purposes. It is an individual-level dataset covering all workers liable to social security. The data originate from the mandatory German notification procedure for social security, which compels all employers to keep the social security agencies informed about their employees. Some spells in the dataset

[1] A detailed description of the SIAB dataset is provided in Ganzer et al. (2017).

[2] Caliendo and Uhlendorff (2008) find that only 3% of the unemployed workers and only 1% of the employed workers enter the state of self-employment in Germany annually.

suffer from overlapping notifications that can occur when data are merged from different administrative sources. For instance, a worker who receives unemployment benefits may have a part-time job, or an employed worker can lose his second job and become part-time employed. Further, the dataset has a structural break in 2005/2006 after the labour market ("Hartz") reforms had been implemented. From that point on, workers receiving unemployment aid were no longer reported. To avoid inconsistent observations, we only consider spells where the employment status is either "employee liable to social security contributions" (defined as: employed) or "receives unemployment benefits" (defined as: unemployed).

The main advantages of the SIAB dataset are the high data quality and its fine regional resolution. For each spell, employed workers are assigned to one of 328 different regions. The regions mostly correspond to "Landkreise" (NUTS 3), some regions have been merged by the data provider to avoid too small numbers of observations (see the appendix of Antoni et al., 2019). There is no regional information for the unemployed. We impute their region as follows: If unemployed workers were previously employed, we assume that they still live in the region of their last job. For unemployed workers who have never had a job before, we assign the region of their next observable job. Workers without any regional information at all are irrelevant for our study and are excluded.[3] Finally, since data for East Germany are nonexistent before the reunification and rather unreliable until 1995, we focus on the years 1995 until 2017. Our final dataset includes more than 30 million spells.

The wages reported in the SIAB are nominal. In order to derive regional real wages, one needs a regional price index. Weinand and von Auer (2020) is the first study that calculated regional prices based on official data of the Federal Statistical Office. Their methodology is data driven and does not need any restrictive assumptions on the regional patterns of prices. The regional price index is, however, only available for May 2016. A dynamic perspective on real regional wages is unfortunately not yet possible.

[3]This reduces the number of spells by about 4%.

3 Aggregated stylized facts

In this section, we present stylized facts about labour market quantities in Germany without a fine-grained regional disaggregation. We consider aggregated unemployment rates, earnings, transition rates between employment and unemployment as well as migration rates for all German regions together.

3.1 Unemployment and earnings

We proceed to compare the monthly unemployment rate implied by the SIAB dataset to the unemployment rate reported in official statistics. To calculate the unemployment rate, we divide the number of unemployed workers by the number of all workers in the SIAB sample *on the first day of each month*. This unemployment rate definition obviously results in a lower rate than the official statistics (see figure 1, top panel). However, the difference between the two rates is stable over time which means that the historical developments are mirrored (e.g. high unemployment in 1997/1998 and around 2005, a short peak after the financial crisis in 2009/2010).

The bottom panel of figure 1 depicts the development of mean real earnings. Nominal earnings have been inflation adjusted by the nation-wide CPI reported by the Federal Statistical Office. The overall mean (black line) declines until 1998, rises slowly up to about 2010 and increases more strongly since then. There is no notable drop during the financial crises. We classify the regions into three equally large categories: rural, suburban and urban regions.[4] The mean real earnings of the three categories are shown by the coloured lines in figure 1. Evidently, earnings are lowest in rural regions and highest in urban regions. Suburbian earnings are close to the overall average. The gap between the three categories is roughly constant over time. There are large differences in the level of daily wages. Workers in rural regions earn on average 20 Euro less per day than workers in urban regions.

[4]The three groups are distinguished by their population density. If the number of inhabitants per square kilometre is 157.88 or less, the region is categorized as rural. If the density is between 157.88 and 367.01 it is a suburban region. Regions with densities of more than 367.01 are defined as urban.

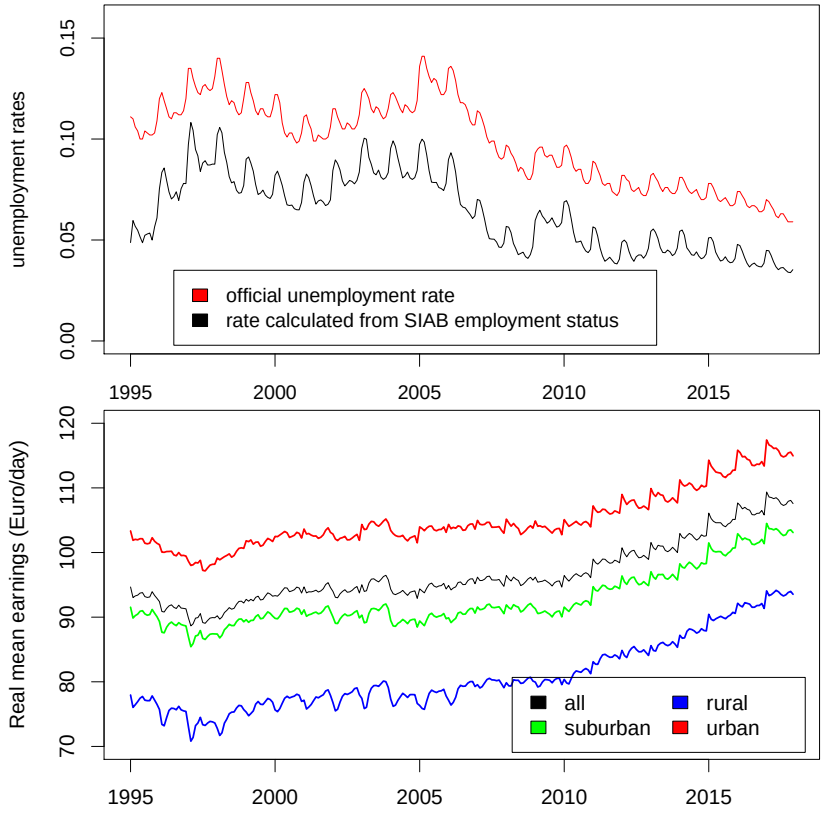

Figure 1: Top panel: Unemployment rates. The red line shows the official unemployment rates, the black lines shows the unemployment rates calculated from the SIAB data as described in the text. Bottom panel: Mean real earnings for urban (red), suburban (green) and rural (blue) areas, the overall mean real earnings are shown by the black line.

3.2 Employment status transition rates

There are three relevant employment status transitions: An unemployed worker can find a job and enter the employment status (job finding transition rate, UE); an employed worker can leave his current job for a new job (job to job rate, EE); an employed worker can enter the unemployment pool (separation rate, EU).

We compute the transition rates by dividing the number of transitions in month t by the number of workers in the origin status in month $t-1$. A transition is defined as a change in the status from the first day of month $t-1$ to the first day of month t. This

definition ignores transitions within months. For example, an employment relationships that began after the first of a month and did not last until the next month, is not counted. Nordmeier (2012) has shown that a monthly measurement underestimates the number of transitions by roughly 10%. However, as it is common to use monthly estimates in most studies, we follow the literature. Figure 2 displays seasonally adjusted worker flows[5] during the sample period from 1995 to 2017.

The monthly UE flow rate varies greatly over this period around an average of 8.06%. At the beginning of the sample period it drops from nearly 10% by more than 4 percentage points and in some months it even falls below 5%. After the labour market ("Hartz") reforms[6] in 2005, the UE rate increases again to over 10% with a small trough in 2009 during the world financial crisis.

The monthly job to job transition rate EE has a mean of 0.88%. While the time series is always close to its mean, it exhibits several positive outliers in the beginning of the observation period and several negative outliers before the structural break in 2005. Taking into account that the average tenure of young workers (see Rhein and Stüber, 2014) became shorter over the last decades, it is surprising that we cannot observe an increasing trend for the EE rate.

The separation rate EU is quite low over the entire time period. About 0.59% of workers transit from their job into unemployment each month. The EU time series shows that there were three recession periods with relatively high separation rates, namely around 1997, 2005 and 2009.

Business cycle fluctuations are associated with large swings in the labour market conditions (see Shimer, 2005). The larger probability to lose a job and the smaller probability to find a job in a recession is obviously one of the major costs of an economic

[5]Because of a change in the reporting system, we computed the seasonal components separately for the years before and after 2005.

[6]The German government restructured the federal employment agency to enhance the matching process of unemployed workers to jobs. Especially, the Hartz IV reform reduced unemployment benefits substantially and abolished long-term unemployment benefits. For more details see Hartung, Jung and Kuhn (2018).

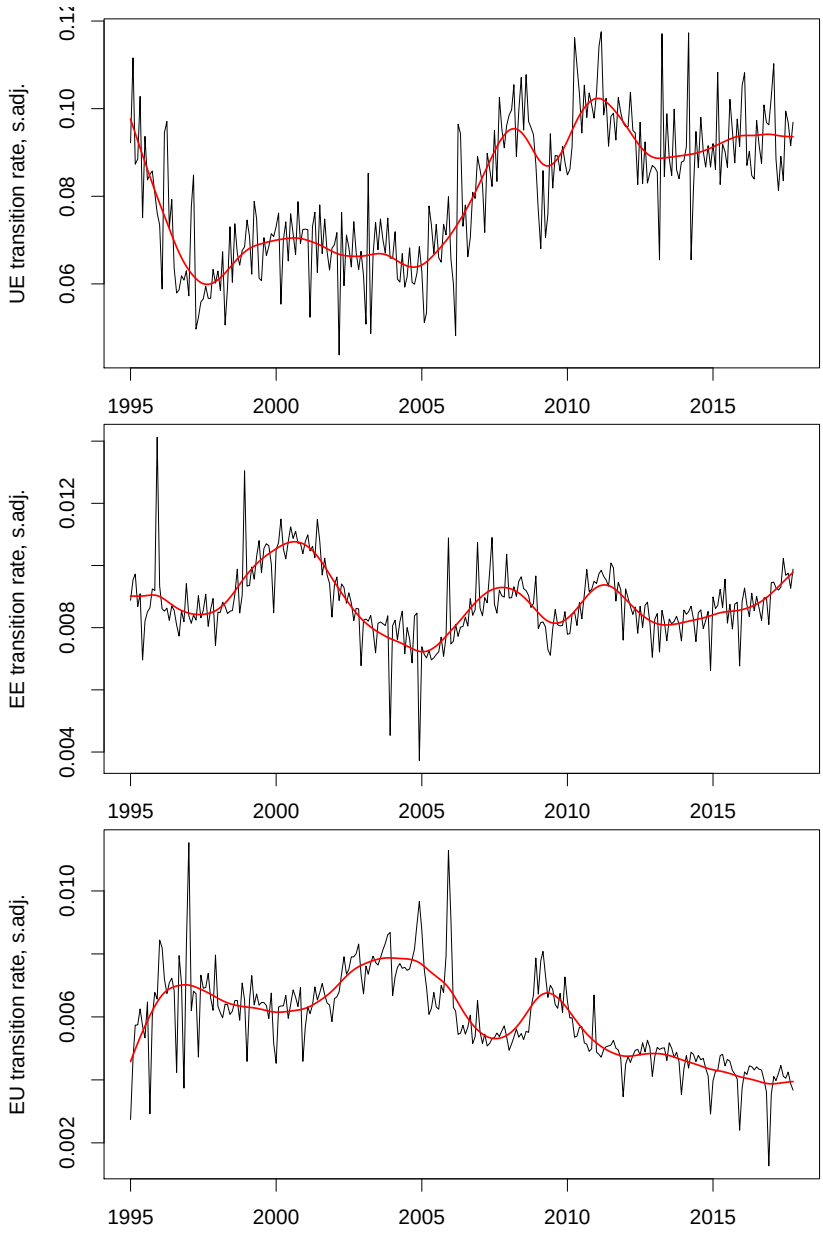

Figure 2: Employment status transition rates. Top panel: unemployment to employment transition rate UE; middle panel: job to job transition rate EE; bottom panel: employment to unemployment separation rate EU; all series are seasonally adjusted, the red lines are smoothing splines.

downturn. Contrary to common expectations our time series do not indicate that the business cycle volatility is lower in Germany than in the U.S. A comparison of both countries (see Shimer, 2005 and Fujita and Nakajima, 2016) shows that business cycle volatility in the German labour market is approximately twice as large as in the U.S. At the same time the level of worker flows is substantially lower[7] because German firms tend to rely on longer-term relationships with their employees. When the economy is hit by a persistent shock, this has a larger effect on firms' present values under long term relationships. If workers enjoy high job protection, expected future productivity is relevant, i.e. the higher the level of job protection, the more the present value of a firm is affected.

3.3 Migration rates

We now consider aggregate measures of workers' movements across regions within Germany. We define a worker as a mover if his or her region of work at the first of month t differs from the first of month $t-1$. Figure 3, top panel, shows the aggregated time series of moving rates from 1995 to 2017. On average, around 0.56% of workers move to a workplace in another region each month. We observe a steep increase in the beginning of the period, which leads to the maximum of nearly 0.75% at the end of the last century. Then followed a decrease to 0.5% in 2005. During the remaining time period, the proportion of movers fluctuates less around its mean. For employed workers the average share of movers is only 0.42%, see the middle panel in figure 3. It varies with a similar amplitude as the total population over the sample period.

The fraction of unemployed movers, depicted in the bottom panel, is close to 2% until 2005 and then experiences a steep increase up to 4.5% with a small and short trough in 2009. The moving rates decline in times of high unemployment rates (see 1998, 2005 and 2009).

To summarize, the main stylized facts about aggregated labour market quantities are: significant changes after the "Hartz" reforms and the financial crisis; high volatilities in

[7] More details can be found in Menzio and Shi (2011) or Jung and Kuhn (2014).

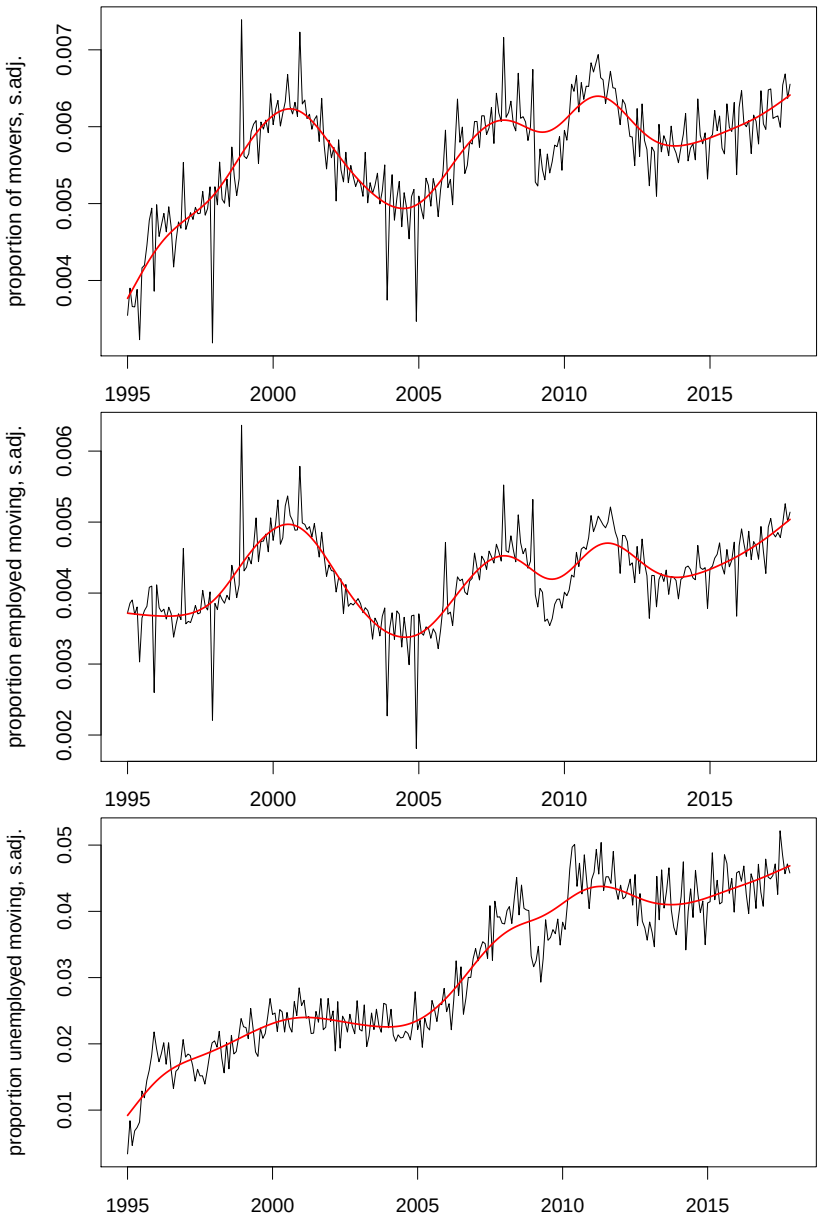

Figure 3: Moving rates. Top panel: proportion of all workers that move in a given month; middle panel: proportion of all employed workers that move; bottom panel: proportion of all unemployed workers that move; all series are seasonally adjusted, the red lines are a smoothing splines.

the transition rates; the job finding rate recovered from an intermediately low; a very remarkable upward trend in the moving rate of unemployed workers.

4 Regional stylized facts

4.1 Regional characteristics

In this section we describe the pattern of regional population densities, price levels (overall and housing) and nominal wages. Information about the population distribution is drawn from the German Federal Agency for Cartography and Geodesy. Germany has only few large cities, and only four of them have a population over 1 million: Berlin, Hamburg, Munich and Cologne. On the other hand, Germany has a large number of medium sized and small cities. In total there are currently 82 cities with a population of more than 100,000 inhabitants. These cities account for approximately 2/3 of the population. However, Figure 4 (top left) also demonstrates that there are agglomeration areas around the urban centres Hamburg, Frankfurt, Stuttgart, Munich and Dortmund. The latter is only the 9th largest city, but its metropolitan area is home to over 5 million people ("Ruhr area"). The northern parts in East Germany as well as large parts in Bavaria are rural, sparsely populated areas. The same holds for the border zones between Thuringia and Hessen, or Brandenburg and Saxony. After the reunification most areas in the former German Democratic Republic experienced a dramatic fall in number of residents during the 1990s. Exceptions are the urban regions of Berlin, Dresden and Leipzig. Currently, 17% of the German population live in East Germany.

Regional difference in nominal wages do not always reflect differences in purchasing power. To this end, regional prices have to be taken into account. While the prices of many products are more or less the same all over Germany, this is not true for one of the most important shares of the consumption basket: housing. The level of housing costs is far higher in urban areas than in rural ones. There are hardly any recent empirical studies that look at regional price differences in Germany. A study that is based on the official price measurements of the federal and state-level statistical offices is Weinand

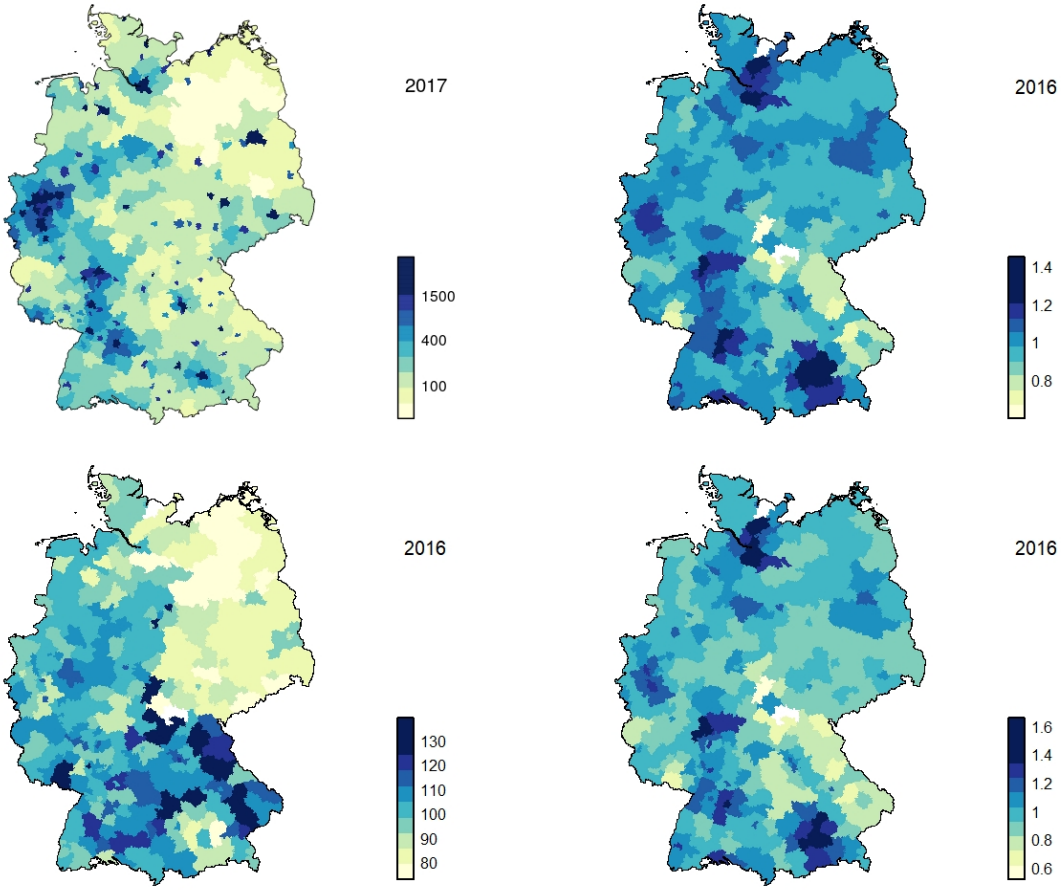

Figure 4: Regional data: population density 2016 (top left), overall price levels 2016 (top right), real wages and benefits 2016 (bottom left), housing price level 2016 (bottom right).

and von Auer (2020). Their price data set[8] from May 2016 includes housing, services and goods, and is normalized by the population weighted average price level, i.e. the population weighted mean is 1. The authors find that price levels are largely driven by housing costs and to a much lesser degree by the prices of goods and services.

The plot of regional price levels indicates that regions with a denser population tend to have higher price levels. Another important factor for local prices are spill-overs from neighbouring countries. It is well known that Luxembourg and Switzerland

[8]Price levels for the regions Plön and Hildburghausen are not available.

have substantially higher price levels than Germany has. To avoid these costs, some expenditures are shifted to the German border regions which leads to a price increase there. Furthermore, it is not surprising that figure 4 also reveals inner-German spillover effects. The price levels tend to be higher in the neighbourhood of expensive regions than in the neighbourhood of inexpensive regions. The positive spatial correlation is mainly driven by housing. Overall, the price level in the most expensive region, Munich, is about 27% higher than in the cheapest region.

4.2 Nominal wages and benefits

Figure 5 shows the level of average daily gross wages or benefits for selected years between 1995 and 2017. Wages exceeding the contribution ceiling for statutory pension insurance are only reported up to this limit. Hence, the data are right censored. The contribution ceiling is time-varying and roughly twice as large as the overall average daily wage. Wage differences are especially large between East and West. Between 1995 and 2017 full-time employees in Germany's western states earn a daily average from 84.8 to 95.9 euros, their colleagues in East Germany earn about 30% less, i.e. from 56.5 to 72.3 euros. While the gap has been narrowing in recent years it is still substantial. Of course individual wages do not only depend on the region of work but also on many other factors such as the employees' experience, their qualification and the industrial sector. For example, two cities with very high wages are Wolfsburg and Ingolstadt (with less than 140,000 inhabitants) which are the headquarters for Volkswagen or Audi. Other high-wage industries in Germany are pharmaceuticals, banking and aviation. As a result we observe high salaries in Munich, Frankfurt and Hamburg. The highest wages at the state level are paid in Hesse and Baden-Württemberg. The annual "wage atlas" of Bierbach which analyses more than 490,000 observations confirms this statement. It reports that wages in Hesse and Baden-Württemberg are 14.1% and 8.6% higher than the nationwide average.[9]

[9] A further aspect in this report is the gender wage gap, which we do not discuss here.

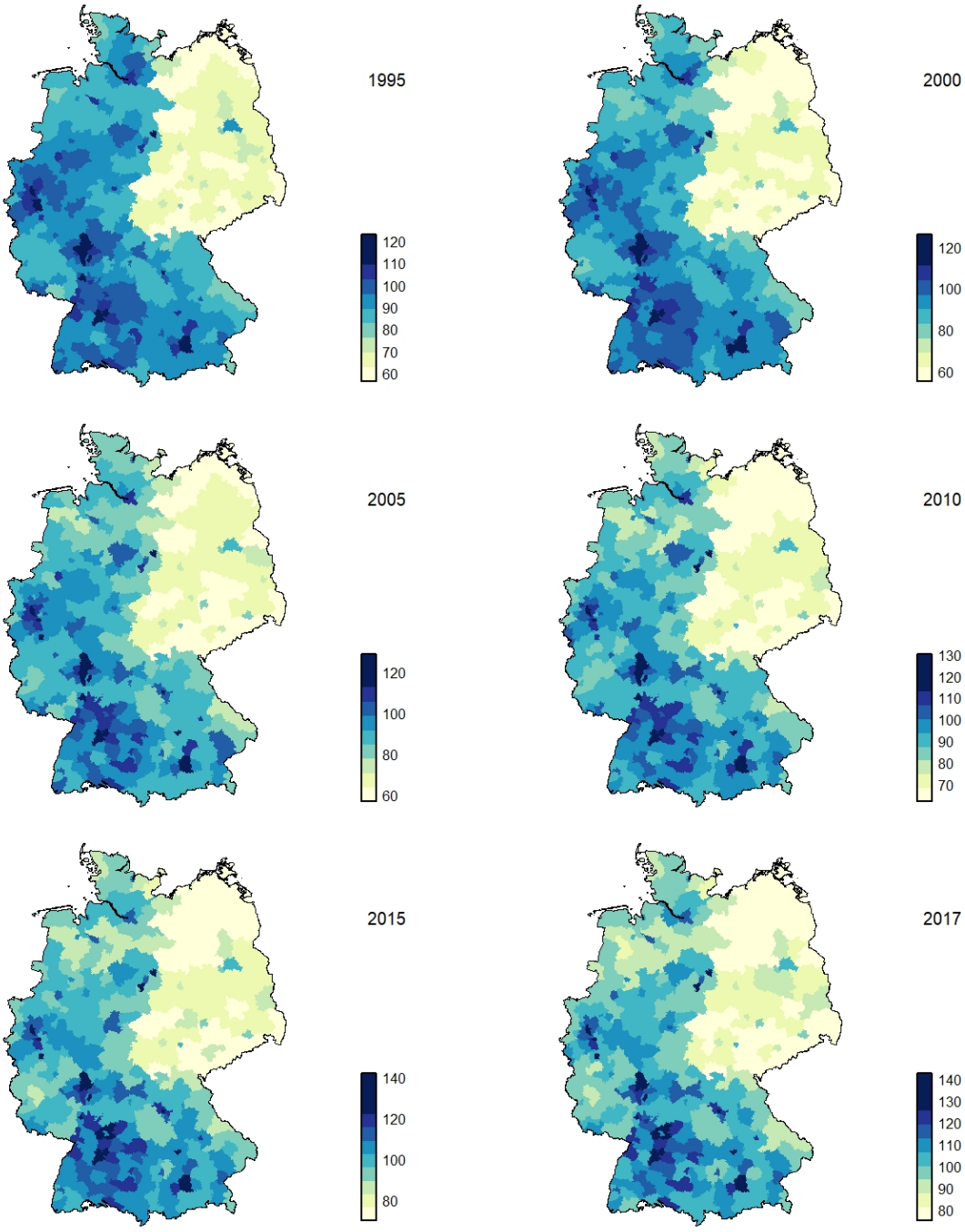

Figure 5: Regional wages and benefits. Note that the colour scale differs between the years.

4.3 Unemployment rates

Concerning unemployment rates (see figure 6), there is substantial regional variation. High unemployment rates persist primarily in the East where the average unemployment rate between 1995 and 2017 amounted to 9.9% with a maximum of nearly 20% in the late 1990s.[10] Medium levels of unemployment rates exist in the northern and central West German Federal States including Schleswig-Holstein, Lower Saxony, North Rhine-Westphalia, Hesse, Rhineland-Palatinate and Saarland. Very low rates are observed in Southern Germany, i.e. Bavaria and Baden-Württemberg where the average unemployment rate is 5.2% with a minimum of 3.3% in Biberach.

Looking at the evolution over time, the unemployment rate has been falling sharply all over Germany. In 2017 no East German region had an unemployment rate of more than 8%, some regions in the West even reached full employment (with unemployment rates of less than 2%). The large and persistent differences between East and West Germany can be considered as a consequence of the long lasting German division. However, the differences in unemployment rates are clearly less persistent than the differences in the wage level. The clear distinction in the unemployment level between East and West that was visible in 1995, is hardly discernible any longer in 2017.

4.4 Migration rates

Figure 7 displays net migration rates for selected years between 1995 and 2017. An advantage of the subdivision into 328 relatively small regions is that only movements within a region remain undetected. A coarser regional division (e.g. by NUTS 1, "Bundesland") would result in much lower rates since most migration occurs over shorter distances (see also section 5). In general, we observe low net migration rates for all regions, with the exception of Munich in 2005. A common finding in the literature is that migration flows from East to West Germany are substantially larger than in the other direction since the reunification. Our calculations confirm a relatively low and decreasing net migration from East to West over time. While our data indicate negative net migration in 2017

[10] These numbers are lower than the official unemployment rates as explained in section 3.1.

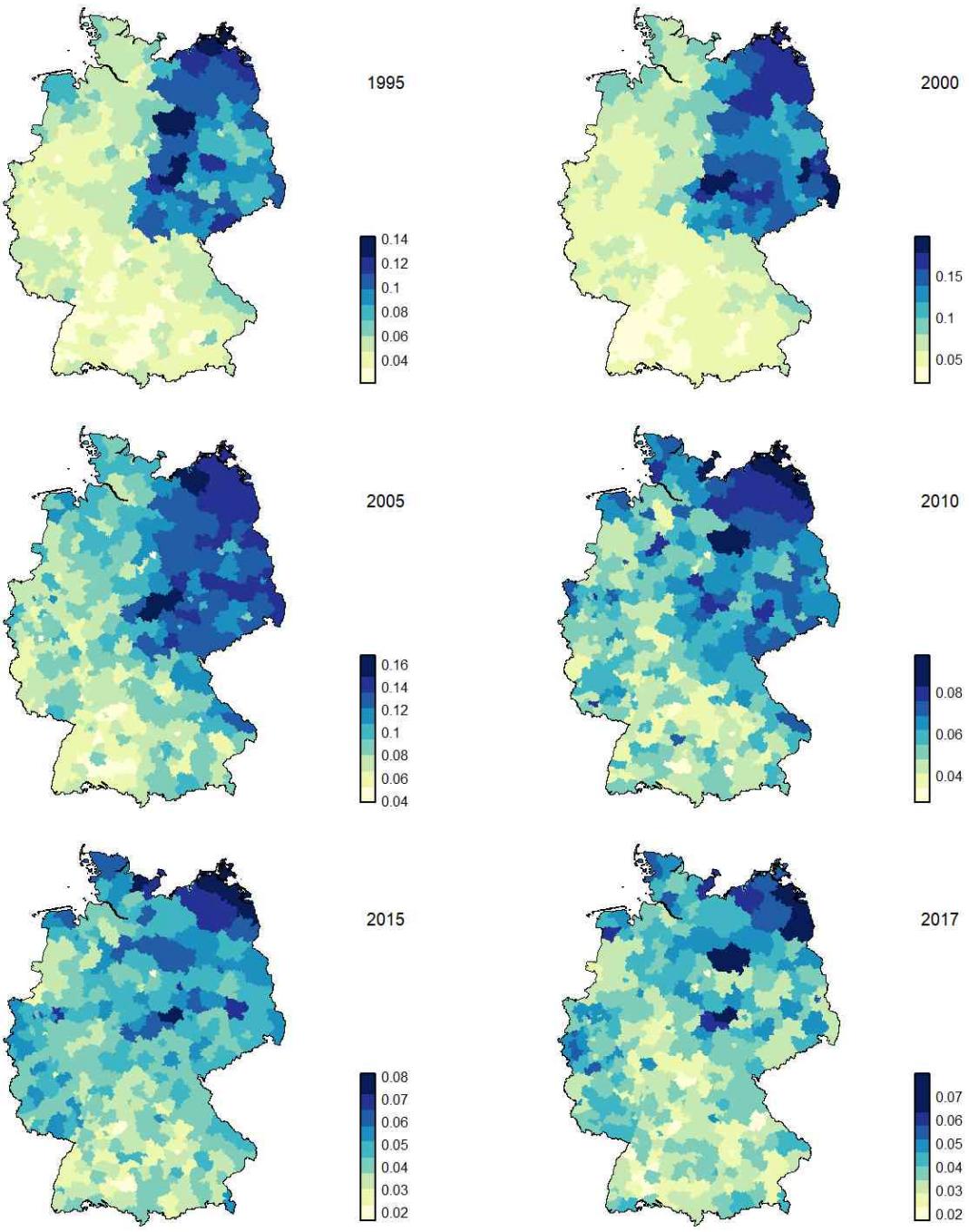

Figure 6: Regional unemployment rates. Note that the colour scale differs between the years.

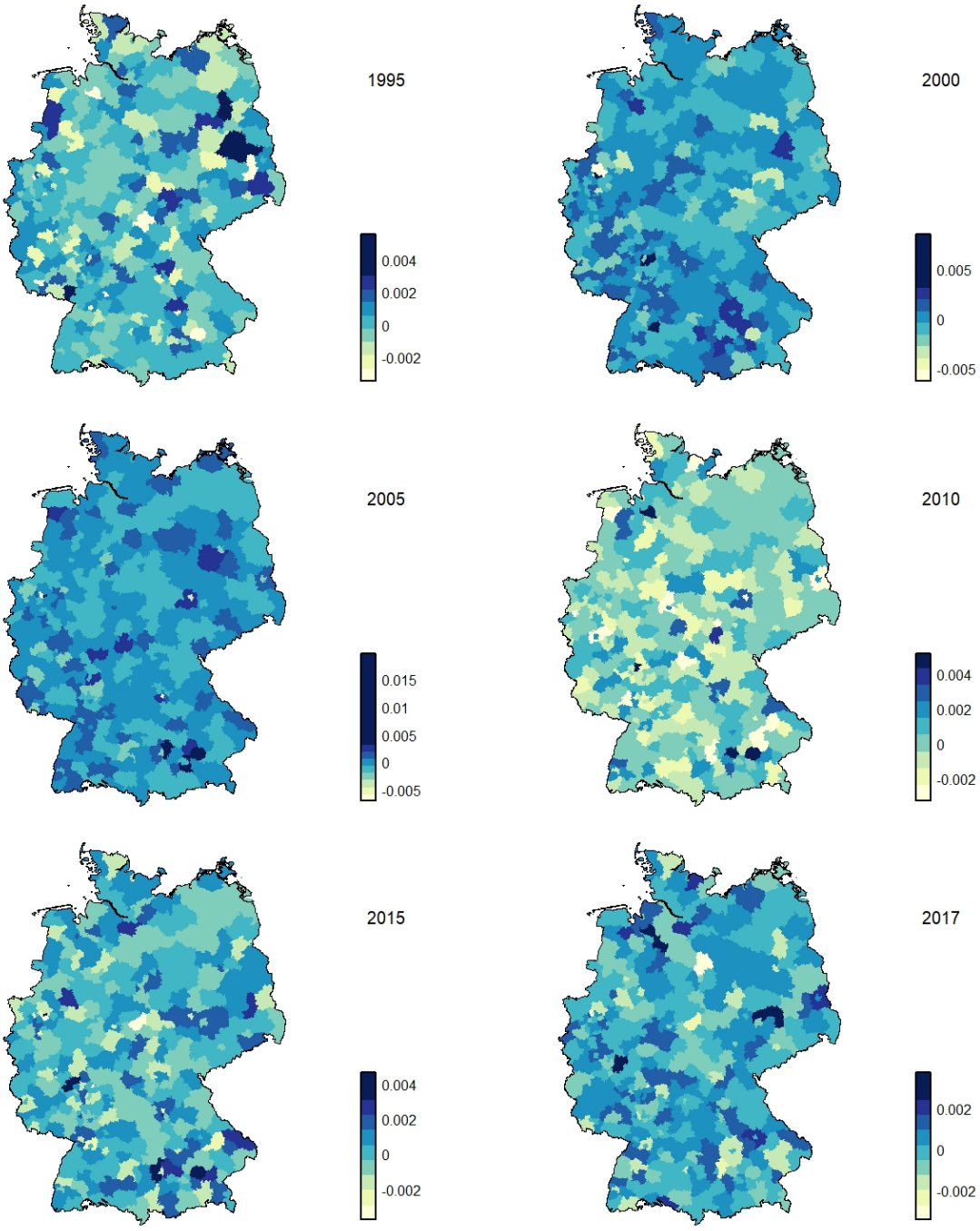

Figure 7: Regional net migration. Note that the colour scale differs between the years.

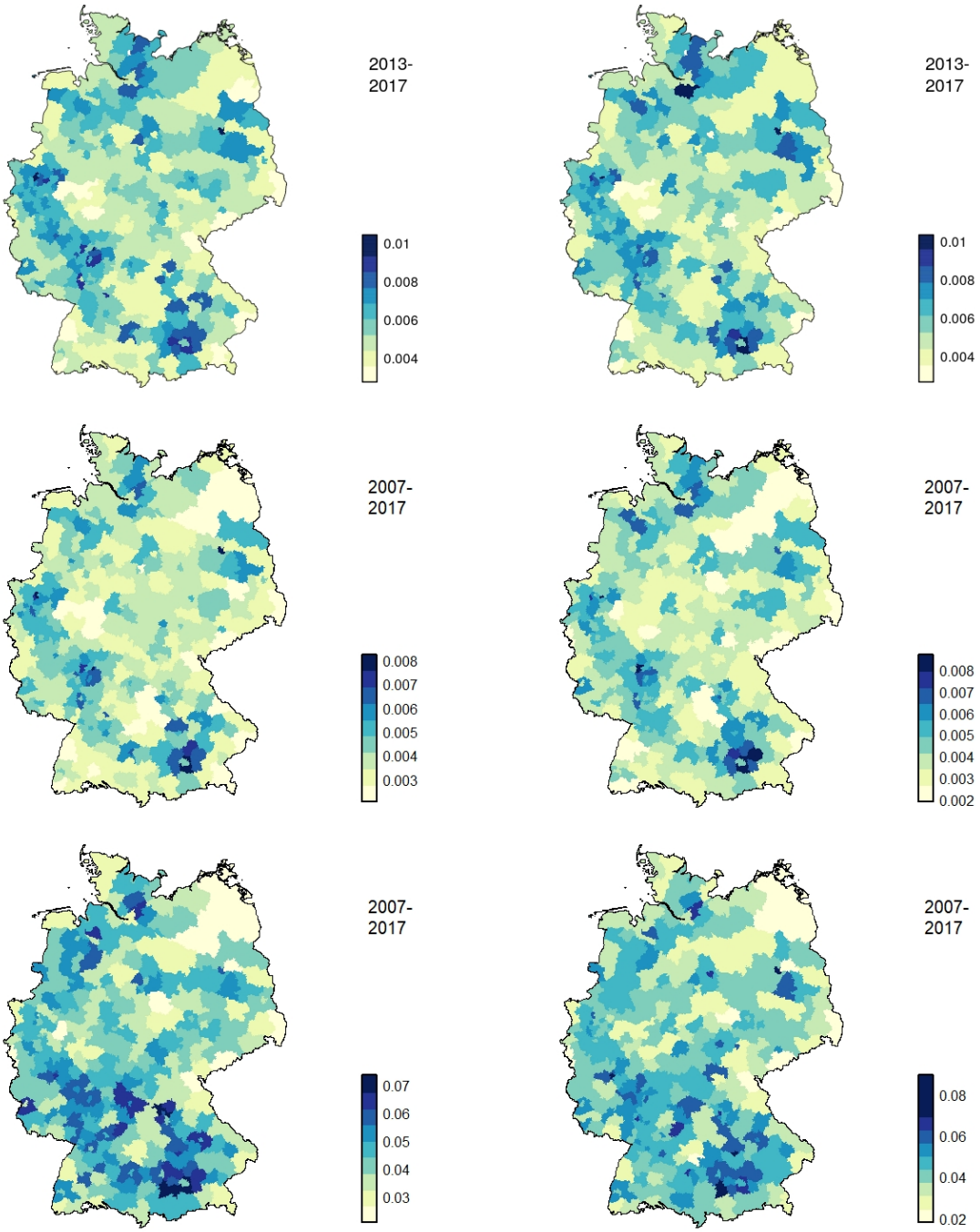

Figure 8: Regional movements (out of/into regions): all workers (top), employed (middle), unemployed (bottom). The rates are averaged over the years 2013-2017.

for East Germany (-0.0034%), Bangel et al. (2019) computed a positive net migration for the first time since reunification.[11]

Further, we plot migration rates between emigrants (left) and immigrants (right) in figure 8, because net values can hide the extent of mobility between regions. The rates are averaged over 2013 to 2017. The regions with the lowest immigration rates are Saxony, Saxony-Anhalt and Mecklenburg-West Pomerania. Areas around metropolises such as Munich, Frankfurt or Hamburg experience the highest immigration rates. This development reflects a recent urbanisation trend. Looking at the differences between employed and unemployed workers, we find that employed workers are less likely to move and their movements are more concentrated into (sub)urban areas.

4.5 Relation between unemployment and other variables

We proceed to take a closer look at the relation between regional unemployment and other variables. Figure 9 shows the relationship between regional unemployment rates and population density; worker flows; movements; and wages (averaged over the years 2013 to 2017). The density tends to be negatively correlated with unemployment rates in sparsely populated regions. More precisely, doubling the density is associated with a decrease of the unemployment rate by nearly 1.5 percentage points for a population density of up to 200 inhabitants per square kilometre. For higher densities this association vanishes. For example, we observe high unemployment rates in the Ruhr area, represented by Dortmund, while there are equally densely populated cities across the country with low unemployment rates.

For employed workers we do not find a connection between moving rates and unemployment, but there is a negative relationship for unemployed workers. This is in conflict with the neoclassical adjustment theory that workers migrate from high to low unemployment regions. In particular, regions in East Germany suffer from this effect. In figure 9, we present Vorpommern-Greifswald as a representative region for the disadvantaged

[11] This is probably caused by a different calculation of the migration rate. We consider the place of work, while Bangel et al. (2019) consider data about the place of residence.

East. Berlin and Munich are located below the regression line, which is probably due to migration between districts within their regions. Positive outliers might be explained by geographical proximity to economically powerful centres. An outstanding example is Erding, a satellite city located 31 km north of Munich. Many workers living in Erding tend to search for a job in Munich.

For the job finding rate (UE) and the job to job rate (EE) the scatter plots do not show any clear pattern, but the separation rate (EU) exhibits an approximately linear relationship to the unemployment rate. This association is presumably amplified by the fact that, in our calculations, workers who lose their job, and unemployed workers, by definition, do not move unless they find a new job.

Finally, we investigate the relationship between unemployment and income and benefits. Since unemployed workers are less likely to move and tend to be less skilled, wages are negatively correlated with the unemployment rate. This effect is supported by the additional bargaining power of firms in case of larger unemployment pools. Taking a closer look at the wages of movers, we observe substantially higher wages for immigrants than for emigrants, but they are still below the region's average wage level.[12]

5 Regression model

The objective of this section is to describe the relation of interregional migration flows and other variables between 1995 and 2017 by regression techniques. Based on all pairwise migration flows, we estimate the role of demographic, geographic and economic factors for regional mobility. As in Bauer et al. (2019), the regression model is

$$\begin{aligned} M_{ij,t} =& \alpha_i^{orig} + \alpha_j^{dest} + \alpha_t^{time} + \beta_1 \ln D_{ij} + \beta_2 \ln P_{i,t} + \beta_3 \ln P_{j,t} \\ &+ \sum_{s=1}^{n} (\gamma_s \ln X_{s,i,t} + \delta_s \ln X_{s,j,t}) + \varepsilon_{ij,t} \end{aligned} \tag{1}$$

[12] On average wages for immigrants are 40.1% higher than for emigrants, while they are 9.6% lower than the average wage in their new working region.

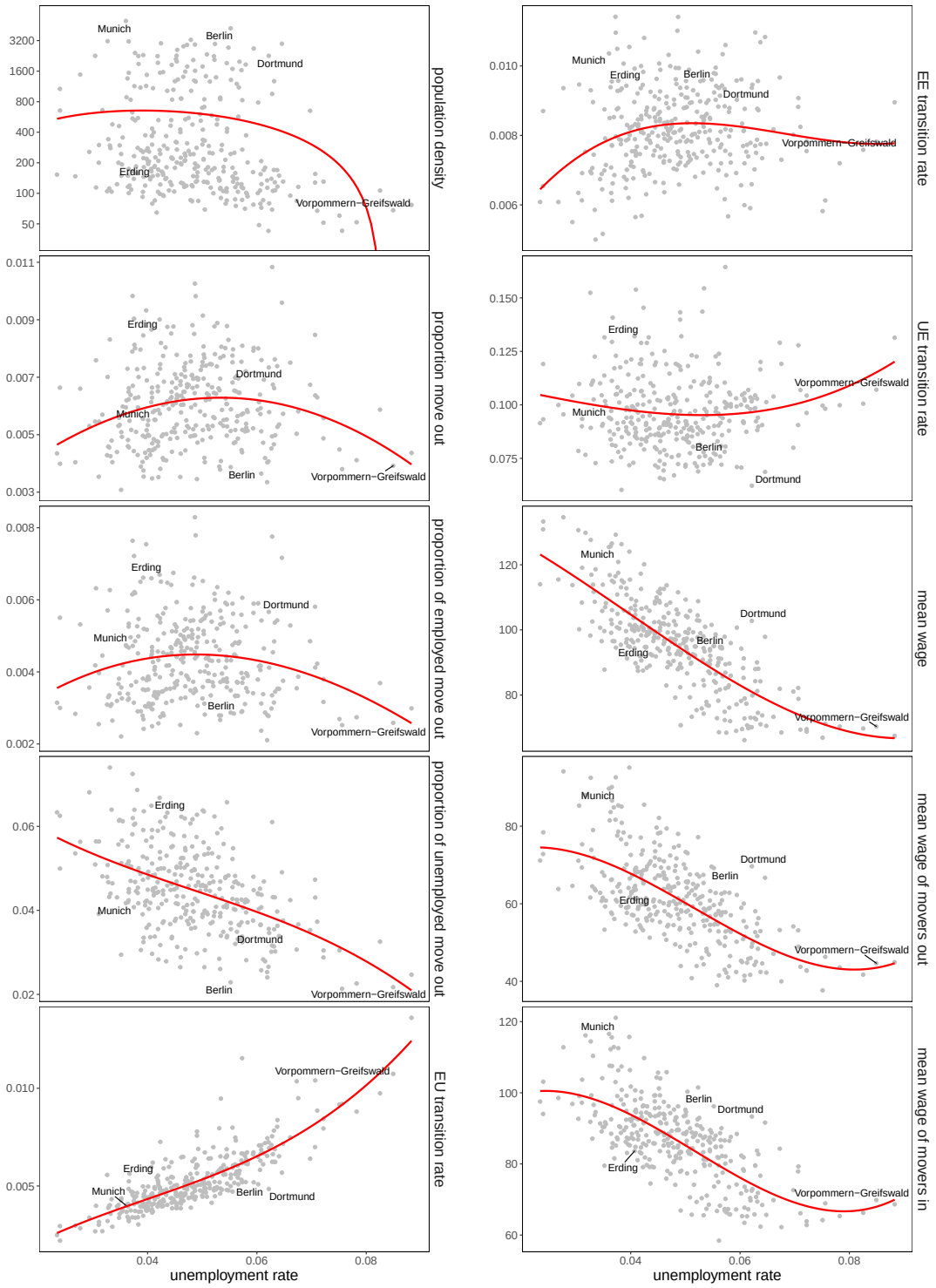

Figure 9: Relationship between unemployment and other economic variables.

where $M_{ij,t}$ is the (absolute) number of migrants from origin region i to destination region j in period t, D_{ij} is the geographical distance between region i and j (calculated as the beeline between the centre points in m), $P_{i,t}$ and $P_{j,t}$ are the population sizes in regions i and j in year t, and $X_{s,i,t}$ and $X_{s,j,t}$ are lists of economic variables in region i and j, respectively, that could act as push or pull factors for interregional migration. Finally, the model controls for regional fixed effects of the origin (α_i^{orig}) and the destination (α_j^{dest}). Time fixed effects are denoted as α_t^{time}.

We estimate (1) using the Poisson pseudo maximum likelihood (PPML) method that Santos Silva and Tenreyro (2006) suggested for estimating gravity models of trade flows. This approach has two advantages. First, it solves the intrinsic problem of log-linearisation that the transformed errors are correlated with the regressors under heteroskedasticity, which leads to inconsistent estimates. Second, taking the logarithm of the dependent variable is not possible for zero observations, hence zero observations need to be excluded from the analysis. In our case the number of observations would have to be reduced by more than 90%. Both these problems are solved by PPML.

The regression estimates are reported in table 1. The distance between regions i and j is a proxy for moving costs and is expected to have a negative sign. The destination population size is expected to affect migration positively, and the effect of population size in the origin region should be negative since a larger population in the origin region implies more job opportunities or better local amenities which tends to reduce workers' incentive to move. Both, distance and population are standard regressors in the literature. As already demonstrated in figures 4, 5 and 6, there is considerable heterogeneity in economic and labour market conditions across Germany's regions.

The effects of most standard variables for our basic model (column 3 in table 1) are in line with extant studies (see Bauer et al., 2019, or Liu, 2018). As expected, the coefficient of distance is highly significantly negative, indicating that a longer distance lowers the number of migrants with an estimated semi-elasticity of -1.81. The unexpected positive coefficient for the population in the region of origin suggests that people living in metropolitan areas are less closely tied to their place of living. The coefficient for the

destination region has almost the same size and confirms the ongoing urbanisation process. Hence, densely populated regions are subject to both higher in- and out-migration. Workers in regions with high unemployment rates are more likely to move to a region with fewer unemployed workers to increase their chances of finding better jobs. This finding does not coincide with the outcome of the scatterplots (figure 9, "proportion of unemployed move out" in the left column) where the proportion of unemployed workers moving out of regions with high unemployment decreases in the unemployment rate. Higher wages in the destination region are associated with lower migration which at first sight is counterintuitive. When we remove the fixed effects or restrict the destination fixed effects to be the additive inverse of the origin fixed effects ($\alpha_i^{orig} = -\alpha_j^{dest}$), our results change substantially (see column 1 and 2).

In the fourth and fifth column, we split the sample according to whether the origin region is in East or West Germany (columns 4 and 5) and compare the different effects of economic and labour market conditions on internal migration in the two subsamples. The estimation result suggests that distance tends to be a more important factor for mobility of workers in East Germany than in the West. Workers in East Germany are less likely to move over long distances. Thus, regional migration especially takes place between larger cities and their hinterland. The values of some coefficients are nearly twice as large as in the West German subsample, indicating that population size and unemployment rates have a greater impact on the incentive to move for workers in East Germany. While interregional mobility of West German workers is less motivated by these factors, wages seem to be have an additional slow-down effect on migration rates in West Germany.

Next, we compare the migration behaviour of men and women (columns 6 and 7). The effects of distance and population size are similar for both genders and coincide with the estimated value of the joint regression (in column 3). Concerning the unemployment rate we observe a notable difference. The migration of men has a stronger link to unemployment than that of women. This is true both for the unemployment rate in the origin region and in the destination region. Wages are only a significant factor to stay

or to leave for men.

Bauer et al. (2019) use the same method to investigate internal migration of different age groups between 2008 and 2014. In contrast to our study they consider the place of residence rather than the place of work. Their analysis of determinants of internal migration is based on data from various sources. Their information on the number of internal migrants is provided by the German population registers. This reporting system is disaggregated to more than 400 regions. Bauer et al. exclude individuals with foreign nationalities as they believe that the behaviour of international migrants might be systematically different from the behaviour of natives. Overall, our results coincide with their estimates. While their effect for population in the destination is not significant, we observe a positive impact. Their outcome suggests that the effect of unemployment in the destination region has a lower effect.

As a concluding remark it is important to keep in mind that all results should be seen as correlations rather than causal effects, since some regressors might suffer from endogeneity. For example, it is obviously conceivable that migration has an effect on unemployment rates and wages, leading to simultaneous causality.

Table 1: Gravity Model of Internal Migration - estimated by PPML; the dependend variable is the (absolute) number of migrants between the origin and destination region.

Variable	**Total**	**Total**	**Total**	**West**	**East**	**Men**	**Women**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
distance	-1.5690***	-0.9533***	-1.8137***	-1,7981***	-2.0934***	-1.7961***	-1.8567***
	(0.0070)	(0.0070)	(0.0038)	(0.0042)	(0.0100)	(0.0040)	(0.0056)
population_o	0.9985***	0.4039***	0.8769***	0.6669***	0.9992***	0.8840***	0.6762***
	(0.0061)	(0.0087)	(0.0529)	(0.0857)	(0.0986)	(0.0520)	(0.0586)
population_d	0.9974***	0.4150***	0.8293***	0.6415***	1.1629***	0.8480***	0.7058***
	(0.0064)	(0.0087)	(0.0554)	(0.0712)	(0.0788)	(0.0541)	(0.0598)
unemployment_o	0.1953***	0.2022***	0.1168***	0.0815***	0.2772***	0.1231***	0.0828***
	(0.0199)	(0.0066)	(0.0249)	(0.0280)	(0.0698)	(0.0206)	(0.0260)
unemployment_d	-0.0764***	0.1388***	-0.2462***	-0.2256***	-0.3535***	-0.1823***	-0.1246***
	(0.0197)	(0.0066)	(0.0252)	(0.0267)	(0.0508)	(0.0203)	(0.0272)
wages_o	-0.0645	0.0407	-0.6810***	-0.2495	0.0859	-0.9185***	-0.2683
	(0.0401)	(0.0338)	(0.1669)	(0.1988)	(0.3996)	(0.1564)	(0.1759)
wages_d	-0.1581***	-0.0312	-0.8830***	-0.6557***	-0.2805	-1.1465***	-0.0600
	(0.0402)	(0.0340)	(0.1652)	(0.1830)	(0.3500)	(0.1548)	(0.1731)
N	2,466,888	2,466,888	2,466,888	2,015,628	451,260	2,466,888	2,466,888
R^2	0.1202	0.6372	0.6879	0.6952	0.6912	0.6462	0.6312
symmetric RFE	no	yes	no	no	no	no	no
non symmetric RFE	no	no	yes	yes	yes	yes	yes
Time FE	no	yes	yes	yes	yes	yes	yes

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

6 Conclusion

This paper provides a comprehensive descriptive picture of Germany's regional labour mobility and highlights substantial dynamics over the period 1995 to 2017. Most extant papers have focused on net migration or state level (NUTS 1) patterns for East and West Germany. Using data from the Institute for Employment Research, our analysis overcomes the limits imposed by state level boundaries that have impeded a fine-grained description of labour migration flows. We investigate interregional migration between 328 (NUTS 3) regions and find that both emigration and immigration mostly take place in more urban and economically prosperous regions. On average, the fraction of movements among unemployed workers is around five times higher than that of employed workers. On the aggregated level, the amount of movements rises and falls procyclicly.

Even 30 years after the fall of the wall there are not only persistently large differences in unemployment rates, population density, wages and education levels between East and West Germany, but there are also notable differences between smaller regions. According to plain economic theory, migration acts as a compensator for such disparities, but our results do not show an adjustment trend. In fact, we find that the opposite seems to hold: for unemployed workers emigration has a clearly negative relationship with the unemployment rate.

Finally, our regression results confirm many findings of the existing literature. Distance and population are the most important regressors for migration. Even if – or simply because – family and friends have a big impact on the decision to migrate, most movements are directed to the next bigger city in the surrounding area. Economic variables such as unemployment, wages and the education level play a much smaller role.

References

[1] Antoni, M., Ganzer, A. and vom Berge, P. (2019), *Sample of Integrated Labour Market Biographies Regional File (SIAB-R) 1975-2017*, FDZ Datenreport 04/2019 EN.

[2] Aragon, Y., Haughton, J., Leconte, E. (2003), *Explaining the pattern of regional employment: The case of the Midi-Pyrénées region*, Papers in Regional Science, 82: 155-174.

[3] Bangel, C., Blickle, P., Erdmann, E., Faigle, P., Loos, A., Stahnke, J., Tröger, J. and Venohr, S. (2019), *The millions who left*, Zeit online, available at https://www.zeit.de/politik/deutschland/2019-05/east-west-exodus-migration-east-germany-demography.

[4] Bauer, T., Rulff, C. and Tamminga, M. (2019), *Berlin calling - Internal migration in Germany*, Ruhr Economic Papers, No. 823, ISBN 978-3-86788-956-8, RWI - Leibniz-Institut für Wirtschaftsforschung, Essen

[5] Bierbach, P. (multiple years), *Gehaltsatlas: Eine empirische Untersuchung zum Zusammenhang von Region und Gehalt in Deutschland*, Annual series of Gehalt.de.

[6] Caliendo, M. and Uhlendorff, A. (2008), *Self-employment dynamics, state dependence and cross-mobility patterns*, IZA Discussion Papers 3900.

[7] Fujita, S. and Nakajima M. (2016), *Worker flows and job flows: A quantitative investigation*, Review of Economic Dynamics, 22: 1-20.

[8] Glorius, B. (2010), *Go west: Internal migration in Germany after reunification*, Belgeo (Online) 3, doi: 10.4000/belgeo.6470.

[9] Hartung, B., Jung, P., Kuhn, M. (2018), *What hides behind the German labor market miracle? Unemployment insurance reforms and labor market dynamics*, IZA Discussion paper 12001.

[10] Heiland, F. (2004), *Trends in East-West German Migration from 1989 to 2002*, Demographic Research, 11: 173-194.

[11] Hunt, J. (2000), *Why do people still live in East Germany?*, NBER working papers 7645.

[12] Jung, P. and Kuhn, M. (2014), *Labour market institutions and worker flows: Comparing Germany and the US*, Economic Journal, 124: 1317-1342.

[13] Liu, L. Q. (2018), *Regional Labor Mobility in Spain*, IMF Working Papers 18/282, International Monetary Fund.

[14] Menzio, G. and Shi, S. (2011), *Efficient search on the job and the business cycle*, Journal of Political Economy, 119: 468-510.

[15] Molho, I. (1995), *Spatial autocorrelation in British unemployment*, Journal of Regional Science, 35: 641-658.

[16] Nordmeier, D. (2012), *Worker flows in Germany: Inspecting the time aggregation bias*, Labour Economics, 28: 70-83.

[17] Rhein, T. and Stüber, H. (2014), *Beschäftigungsdauer im Zeitvergleich: Bei Jüngeren ist die Stabilität der Beschäftigung gesunken*, IAB-Kurzbericht 03/2014.

[18] Santos Silva, J.M.C and Tenreyro, S. (2006) *The Log of Gravity*, The Review of Economics and Statistics, 88 (4), 641–658

[19] Shimer, R. (2005), *The cyclical behavior of equilibrium unemployment and vacancies*, American Economic Review, 95: 25-49.

[20] Smolny, W. and Peukert, C. (2011), *Interregional migration in Germany: Characteristics and effects for regions and migrants*, doi: 10.2139/ssrn.1782013.

[21] Weinand, S. and von Auer, L. (2020), *Anatomy of regional price differentials: Evidence from micro-price data*, Spatial Economic Analysis, doi: 10.1080/17421772.2020.1729998.