

Kremer, Anna

Working Paper

Home is where the heart is? How regional identity hinders internal migration in Germany

CEPIE Working Paper, No. 05/20

Provided in Cooperation with:

Technische Universität Dresden, Faculty of Business and Economics

Suggested Citation: Kremer, Anna (2020) : Home is where the heart is? How regional identity hinders internal migration in Germany, CEPIE Working Paper, No. 05/20, Technische Universität Dresden, Center of Public and International Economics (CEPIE), Dresden, <https://nbn-resolving.de/urn:nbn:de:bsz:14-qucosa2-721720>

This Version is available at:

<https://hdl.handle.net/10419/224736>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

CEPIE Working Paper No. 05/20

Center of Public and International Economics

**HOME IS WHERE THE HEART IS?
HOW REGIONAL IDENTITY HINDERS
INTERNAL MIGRATION IN GERMANY**

September 2020

Anna Kremer

Editors: Faculty of Business and Economics, Technische Universität Dresden.

This paper is published on the [Open Access Repository Qucosa](#).

The complete Working Paper Series can be found at the [CEPIE Homepage](#) | [EconStor](#) | [RePEc](#)

Home is where the heart is? How regional identity hinders internal migration in Germany

Anna Kremer*

TU Dresden and ifo Institute Dresden

September 2020

Abstract

People are emotional about places. I study the effect of regional identity (“at home”) on internal migration flows in Germany between 1995 and 2017. Regional identity is proxied by measuring how NUTS3 regions were historically affiliated in the former patchwork of Germany. When controlling for the influence of distance, culture (measured by dialects) and regional characteristics, I confirm that regional identity drives migration patterns additionally. Employing the separation effect by the German wall affirms that not only earlier migration or family bonds determine movements instead of regional identity.

JEL classification: R; R23; Z10; J61

Keywords: Internal migration, Regional identity, Historical belonging, Gravity model, Germany.

*Corresponding author: TU Dresden and ifo Institute Dresden, Einsteinstr. 3, 01069 Dresden (Germany), E-mail: kremer@ifo.de

Introduction

Feeling at home is important for all of us. Yet, about 3% of the population in Germany moves across district borders each year (Stawarz and Sander, 2020). I look at whether these movements are driven by where people feel at home.

Migration flows determine, among other factors, whether a place prospers or is left behind. This development of regions is theorised in the core-periphery-model of spatial economics that explains how agglomerations form over space (Krugman, 1991a and 1991b). An efficient allocation of economic activity in the model stems from perfect mobility of workers and them maximising their wages over locations. However, migration is far from economically perfect, among other factors due to a lack of information, uncertainty (Molho, 2013) or non-pecuniary (social) costs (Newbold, 2019). Also, culture plays a role in migration decisions. This is already evident for international migration being limited by cultural differences (Belot and Edrveen, 2012; Adserà and Pytliková, 2015; Adserà, 2015). The work of Falck et al. (2012) and Falck et al. (2018) shows that culture is also a determining factor for internal migration in Germany. The authors use linguistic differences as a measure for cultural dissimilarity and estimate how these hinder migration (Falck et al., 2012) and by what wage premium moves to a culturally distant place are compensated (Falck et al., 2018). Similar to culture, feeling connected to a place is relevant for moving or stay-at-home decisions. That people feel at home in one place but not naturally in another therefore updates economic theory. This reasoning, that feeling at home in a place affects internal migration, is analogue to expecting that international trade or migration is hindered by country borders (e.g. McCallum, 1995; Cushing and Poot, 2004 or Helliwell, 1997), when not entailing only institutional but also emotional difficulties in crossing (invisible) borders.

The German term “Heimat” expresses connectedness to a place and encapsulates how places shape people. “Heimat” is approximately translated as “home(land)” and entails in its different understandings attachment to a place, origin and memories, emotions, social contacts, culture and traditions (Sturm, 2016, p. 82). Next to the subjective experiences that form “Heimat” individually, it is also shaped by history. According to Mitzscherlich (1997, p. 227) this gets incorporated as a foundation for self and self-determination. Likewise, the concept of place identity, a more global idea, refers to having a personal connection to a place. It is considered as being a sub-structure of personal identity and driven by spatial experiences and situations (Proshansky et al., 1983). It is based on definitions of spaces (e.g. categorizations of regions) and on the personal attachment to them (Peißker-Meyer, 2002). Moreover, is place identity multidimensional in referring to different levels (local, regional, national, supra-national) at the same time (Sturm, 2016, p. 80).

Building on the concepts of “Heimat” and place identity, I hypothesise, that individuals feel belonging to places, resulting from physical personal experiences, social ties and culture, which then affects migration decisions. People are emotional about places, reflected by “Heimat”-feelings, and include spatial categorizations into their identity. I expect, that for example a person from Dresden feels belonging to the region of Saxony and thus rather moves within the region than out of it. Studies that regard mobility of graduates confirm this importance of “Heimat” and being at home: Oggenfuss and Wolter (2019) look at university graduates from Switzerland. They find that about half of the students that have left their home region to study return thereafter. Buenstorf

et al. (2016) similarly regard location choices after graduation in Germany. They confirm the importance of moving back home (or staying in the university region) but also show that graduates are more likely to move to a region similar in dialect and settlement structure to their home region. Kaplan et al. (2016) find in a survey that students in Dresden are more likely to intend to stay in Saxony after graduation if their are from that region.

Following a similar idea, Mühler and Opp (2004) and Sedlacek et al. (2009) both find that regional identity affects migration adversely. Their studies are case studies, based on a small sample of survey participants in West Saxony (Mühler and Opp, 2004) or an Austrian village (Sedlacek et al., 2009). Since the analysis has been limited to small samples or subgroups so far, I use a more integral approach to study the effect of belonging to a region on migration. For this I extend the work of Falck et al. (2012) who look at how culture (proxied by dialects) affects internal migration in Germany. The authors establish a connection between a measure of dialect dissimilarities, which is continuous over space, and internal migration between 2000-2006. I prove an additional impact of regional, historically reasoned, borders that potentially links to a place identity. I thus add to the literature in estimating the correlation between (today's invisible) regional borders that affect inhabitants and intra-country moving decision. I connect this effect to a regional identity, because the concept of "Heimat"/place identity and anecdotal evidence show that people refer to whether they feel at home in a region when they are moving.

In the first section I now turn to presenting evidence on the relevance of place identity and to how I proxy for regional identity. I then introduce the migration gravity model and subsequently comment on the data and estimation strategy. Next, I explain the empirical outcomes and make some robustness checks. Finally I conclude my findings.

Regions and identity

The idea that places play a role in shaping our identity ("Heimat") seems intuitive when thinking about new acquaintances and the fact that origin and/or current place of residence is usually mentioned in an introduction of oneself. Accordingly several scholars confirm that individuals and groups are formed by their connection to places: The concept of place identity originates from psychology (Proshansky et al., 1983) and is also investigated empirically in this discipline. With this background, Plaut et al. (2012) confirm that places matter for the self by comparing Boston and San Francisco. They connect the reciprocal influence between individual selves and spatial traditions to differences in personalities that have been dismissed as stereotypes so far. In another study, Anton and Lawrence (2014) in Australian bush-fire risk areas find that for inhabitants of rural areas local place identity is higher than in urbanised areas. Ríos and Moreno-Jiménez (2012) compare natives and immigrants in Málaga, Spain and find higher place identity among natives on several levels (neighbourhood, city and country). Immigrants identify more with a place when they live there for longer or are employed and own their home. With a geographical background, Paasi (2002) proves the existence of place identity in Finland by the pattern of intra-country migration. Besides psychology and geography as natural forerunners in the research on place identity, the concept is also incorporated in environmental studies (Hernández

Bernardo et al., 2010), social sciences, including especially sociology and political science (Mühler and Opp, 2004; Sedlacek et al., 2009) as well as history and planning (Raagmaa, 2002) or marketing (Panzone et al., 2016). Notably since 2005 the topic has been receiving a growing interest with a surge of publications throughout the last decade. A good overview is found in Peng et al. (2020).

In Germany, “Heimat” has a meaning close to regional identity. How much people feel connected to their place is investigated in several surveys: The reference category of the surveys is mostly the federal states, even though Sturm (2016) suspects that these are often too large and conglomerated to feel attached to. Indeed, the Bertelsmann-Stiftung (2008) shows that identification with the federal state is no higher than 25% for the inhabitants anywhere. They conducted a survey among 4000 participants by telephone. People living in Bavaria, Brandenburg, Bremen, Mecklenburg-Western Pomerania, Saxony, Schleswig-Holstein or Thuringia identify more with their federal state, people in North Rhine-Westfalia, Rhineland-Palatinate and Lower Saxony less. For Bavaria, a separate survey by the Hanns-Seidel-Stiftung (2009) reveals substantial pride in living in the federal state. The study furthermore confirms an attachment to (historic) sub-regions of Bavaria. Other studies show that smaller regional units than federal states seem more relevant for identity formation: At a more local level, Beaman (2020) refers to a “dialect identity” in her study of Swabian dialect speakers and both the European Value Survey (see Förtsch et al. 2019) and the survey of the Bertelsmann-Stiftung (2008), find a high importance of the local municipality for place identity or attachment. The study by the Bertelsmann-Stiftung (2008) also asks participants for their main reason to live at their current place of residence. People predominantly respond that it is their place of birth.

In summary, evidence for feeling at home affecting people’s decisions is found in different studies. Looking at Germany, they show that place identity exists but attachment to federal states is lower than to more local units. The whole extent of a connection to “Heimat” that people (unconsciously) feel, is potentially not reflected by the stated identification in surveys, however.



Figure 1: Germany in 1648, in 1815 and today, source: Diercke Weltatlas (2009)

In order to proxy for connectedness to “Heimat” and regional identity in Germany I make use of the historical landscape. In this I assume that identification with places is higher where regions are historically more congruent and belonged together for a longer period of time. Past times, Germany was a vast patchwork of political entities, that additionally changed throughout time (see Figure 1). It was only in 1871 that

the German empire was founded and all territories were ruled together. Hence, regions are distinguishable until today and traditions vary across Germany (compare Wehling, 2013). Empirical evidence on regional differences in values stems from the European Value Study (e.g. Kaasa et al., 2016). I therefore make use of former borders to proxy for place identity, which is also in line with Paasi (2002) stating that “identity and boundaries are different sides of the same coin” (p.139).

Using the historical borders, congruency and connectedness of regions as proxying for regional identity is supported by data in the aforementioned surveys: Bavaria with higher identification has been historically a relatively consistent region since the Congress of Vienna in 1815, as opposed to the “artefacts” North Rhine-Westfalia or Rhineland-Palatinate from after the second world war with lower identification. Similarly, regional stereotypes are often motivated by the historically different backgrounds. An example are stereotypes of Swabians. Swabia is a historic region of Germany which is not a political entity any more but forms part of Baden-Wuerttemberg. I thus expect that belonging to older, handed down regions is relevant for place identity.

In the next section I turn to presenting the theoretical model.

The migration gravity model

I employ the gravity model developed in Anderson (2011), that translates the trade gravity model from Anderson and Van Wincoop (2003) to the migration context.¹

The model is based on an individual utility maximisation with respect to migration. Individuals migrate if local utility w (wages in the original paper) elsewhere is larger than the domestic one, including the costs of migration, modelled as ice-berg costs $\delta_{ij} > 1$. The utility of a specific place w_i or w_j consists of wages, but also entails other (dis)amenities. Costs refer to all frictions for moving between two places, which are not only monetary costs but also cultural dissimilarity and more. The utility for a migration decision is assumed as logarithmic, entailing declining marginal utility from differences in w . With an idiosyncratic utility part ε , due to unobserved utility but also perception errors, the migration decision becomes stochastic.

$$\exp(u_{ij}) = \frac{w_j}{w_i * \delta_{ij}} * \varepsilon$$

The utility does not only determine for the individual whether (not) to move, but also where to. When looking at the aggregated population (and assuming identical agents except for ε), the individual probability of migration from i to j transforms into the share of migrants from i choosing a particular destination j .

According to this reasoning, flows M from i to j are dependent on population N in i and on the proportional utility from moving to j vs to other destinations k . With

$$G(u_{ij}) = \frac{\exp(u_{ij})}{\sum_k \exp(u_{ik})}$$

being the probability of an individual from i choosing destination j over all others, the

¹The full derivation is found in the original paper.

migration equation becomes

$$M_{ij} = G(u_{ij})N_i = \frac{\exp(u_{ij})}{\sum_k \exp(u_{ik})} N_i = \frac{w_j/\delta_{ij}}{\sum_k w_k/\delta_{ik}} N_i$$

The outward migration friction, proxying for features of a place that make people stay or not, is given by definition as $W_i \equiv \sum_k w_k/\delta_{ik}$ and accordingly $\Omega_j \equiv \sum_i \frac{1/\delta_{ij}}{\sum_k w_k/\delta_{ik}} \frac{N_i}{N}$ as inward migration friction is derived. The two sums account for general (dis)attractiveness of regions, the supply and demand determinants. When defining the whole population by $N \equiv \sum_i N_i = \sum_j N_j$ and the sum of all migration flows ending in j (including those that stay in j) as $N_j \equiv \sum_i M_{ij}$, flows depend on origin and destination potential. In a more intuitive manner, migration flows can then be expressed as

$$M_{ij} = \frac{N_i N_j}{N} \frac{1/\delta_{ij}}{W_i \Omega_j}$$

This state of the model suffices as simple framework for the empirical analysis.

In the estimation, migration frictions δ are proxied by physical distance, dialect similarity and historical belonging. The characteristics of origin and destination (W and Ω) are captured by region-time fixed effects. The estimation equation thus becomes

$$\log M_{ijt} = \text{const} - \sum_l \gamma_l * \delta_{lij} + W_{it} + \Omega_{jt} + \varepsilon_{ijt}$$

After developing the model, the next sections introduce the data and empirical strategy.

Data

Data on the historical belonging between 1820 to 2017 is operationalised by myself based on GIS maps of German regions from Kunz and Zipf (2008) and MPIDR and CGG (2011a and 2011b).² A region is defined as a political entity before 1871, the foundation of the German empire, and as the administrative unit between the national and local level (states or federal states) thereafter. As seen in Figure 1, borders have changed from time to time in Germany. These transformations of regions are taken into account by looking at all points in time when borders changed.

Data on internal migration flows in Germany is provided by BBSR Wanderungsstatistik and limited in this analysis to the movements of Germans over the timespan from 1995-2017 between all 401 German NUTS3 regions. The sample is plagued by a selection bias because only realised moves but no considerations of moves are included. Descriptive statistics are found in Table 1.

As a fundamental component in a gravity model, the distance between origin and destination of a flow is relevant (compare Ravenstein, 1885 already; for Germany in

²Maps from 1820, 1826, 1830, 1834, 1848, 1850, 1853, 1863, 1867, 1871, 1890, 1914, 1920, 1922, 1929, 1934, 1935, 1937, 1939, 1941, 1945, 1946, 1947, 1949, 1950, 1952, 1955, 1957, 1961, 1990 capture all major changes.

Table 1: Descriptive statistics of migration flows in Germany between 1995-2017

District pairs	Years	Sum of all moves	Mean of flows	Share of district pairs with positive flows
160 400	23	51 811 375	14.04	0.73

Parikh et al., 2003; Schneider and Kubis, 2009). Data on physical distance between NUTS3 regions is thus used as a control variable. Because not only geographical distance but also cultural aspects play a role, furthermore a dialect similarity index is included, which positively affects migration flows in Falck et al. (2012). The index is obtained from Falck et al. and mirrors dialect differences from the 19th century that are still well-grounded in reality.³

Other data on control variables is taken from the official databases and provided by BBSR Wohnungsmarktbeobachtung (for detailed sources see the Appendix). These control variables are only used in robustness checks as substitutes for region-time fixed effects.

Data on origin and current post codes of (anonymous) individuals is for an additional analysis kindly provided by Gosling and Potter (2016). The observations stem from participants of an internet personality test and therefore over-represent younger and female inhabitants.⁴

Several descriptive statistics show the sensibility of the data. Therefore, I now turn to the empirical strategy.

Empirical strategy

I suspect that regional identity is an influencing factor for (internal) migration. In the estimation of the migration gravity model for Germany I thus add such an explanatory variable and employ the (dis-)continuity of the historic regional pattern of Germany as a proxy. This is based on the assumption, that in regions that did not change over time it is easier to feel belonging to the region, for example because more traditions are passed down. I moreover assume that districts within a historic or federal state shared a bond during this period and inhabitants feel belonging to the same region. With this idea in mind I regard whether two NUTS3 regions belonged together historically and for how long.

I measure the shared historical tradition of two districts by the time they belonged to the same (federal) state. I compute how all NUTS3 regions formed part of different political entities at every moment in time. I then accumulate the time that two regions belonged to the same territory for every pair of regions. To give an example, Munich and Nuremberg belonged both to Bavaria over the whole timespan, thus 197 years. Because some borders crossed today's districts boundaries, I weigh belonging to a territory by the shares of the different areas. For example, the Alb-Donau-Kreis belonged to Wuerttemberg by 20% of its area in 1820 while the other 80% formed part of Bavaria. Therefore I weigh the time length interval of belonging to either territory by these percentages. Two examples of the resulting proxy are shown in Figure 2: The pattern

³Further discussion is found in the original paper.

⁴Further information on this data set is found in Obschonka et al. (2019).

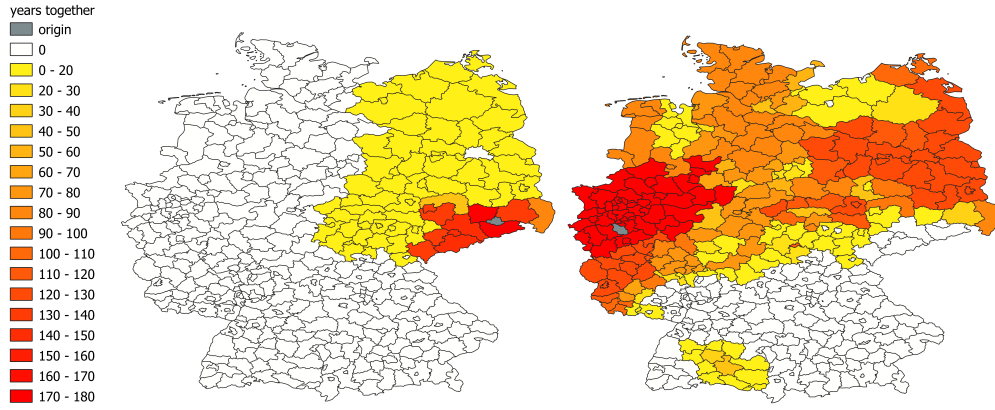


Figure 2: Dresden (left) and Köln (right): Historical belonging, source: based on Kunz and Zipf (2008), MPIDR and CGG (2011a and 2011b)

for Köln is diverse because the region was part of Prussia that enlarged its territory over time. Saxony, where Dresden is located, had rather stable boundaries. Alternative specifications of the measure are mentioned in the sensitivity analysis.

Establishing a causal connection between regional identity and migration behaviour is difficult because they endogeneously shape each other. As opposed to my measure, Abraham and Nisic (2007) for example proxy regional ties by time since last moved. I therefore employ the German wall, which hindered migration for almost 30 years, to find out about the causal relationship. This analysis indicates whether today's migration behaviour still shows systematic differences between historically connected and unconnected regions when earlier migration was largely restricted.

Estimation method

The estimations follow three methodological approaches: Firstly, I apply a Poisson Pseudo Maximum Likelihood method (PPML) to the observations. A PPML is the standard workhorse gravity model approach (Shepherd, 2016) because it proves useful with some observations being zero in the dependent variable (lessening the selection bias to some extent) and provides (non-linear) estimates of the models. The estimator is also consistent for fixed effects and all distributions of data, assumes that the error term is logarithmised and solves heteroscedasticity. The PPML method also serves as base estimation for the robustness checks.

Secondly, I estimate a cross-section of average flows in a similar manner to Falck et al. (2012) with region fixed effects. Yet, I still employ a PPML estimation because the data is heteroscedastic.

As a third method, I make use of a two-step model that estimates the extensive and intensive margin of migration separately. By this it accounts for the selection bias problem that arises because the data only contains realised moves and does not include how many individuals consider moving. The two steps consist of estimating whether individuals decide to move or stay (extensive margin) and how many move where (intensive margin) separately. With this method I follow Egger and Larch (2011) because their approach accounts for heteroscedastic disturbances better than a Heckman sample selection model and makes less restrictive assumptions (Egger and Larch, 2011). I use

origin and destination fixed effects in the Probit estimation of the first step (extensive margin) justified by a Likelihood-ratio-test (LR-test).

In accordance with Egger and Tarlea (2015) I always cluster at the origin-, destination and pair-level. Furthermore I use origin- and also destination-time fixed effects (like Mayda, 2010 or Egger and Larch, 2011) because they control best for multilateral resistances, (dis)attractiveness of districts that are acknowledged by all migrants. Including these effects is also advocated by Olivero and Yotov (2012) for the same reason. I do not make use of pair fixed effects because the variables of interest are then omitted. Also, the estimation of the model only builds on the cross-section dimension of the panel because historical belonging is time-invariant.

In the following I present my results.

Results

I estimate the gravity model by a PPML in two specifications and by a two-step model to account for selection bias. The regression results are interpreted as semi-elasticities due to the log-linear regression form.

The PPML estimation of the baseline model (column (1) of Table 2) shows that if historical belonging is higher by ten years (out of 197 years max) migration flows go up by 7.5%. Likewise, if dialects are more similar by one point (out of 66 max)⁵, movements go up by 6.5%. If distance is 100 km shorter, flows increase by 40.7%. When controlling for socio-economic factors directly⁶ instead of by region-time fixed effects, results are similar (see Table 4).

A model with average flows, estimated as in Falck et al. (2012) but by a PPML, matches in its results with the previous findings (see column (2) of Table 2).

The two-step method of Egger and Larch (2011) further supports the findings. The estimated coefficients for the intensive margin (how many move where, based on all positive flows between districts that make up 73% of the sample) are again very close to the previous findings (see column (3) of Table 2). The first step estimation of the extensive margin (whether people move at all) is also determined by the same variables (see Table 5). However, the explanatory power of the model for the decision to migrate or not is much lower and so are the coefficients.

A graphical illustration of the results from the base PPML is given in Figure 3. It shows how much historical belonging results in higher flows and refers to flows outgoing from Frankfurt a.M. These flows are specified relatively to those without historical ties.

The different estimations only confirm a correlation between historical belonging and migration movements and no causal relation. Earlier migration between historically connected regions potentially induces today's migration and it is likely to factor in for a regional identity. Therefore I now turn to dealing with the endogeneity problem.

⁵The mean dialect similarity of 32 points is e.g. given for Munich and Berlin. A one point higher dialect similarity of 33 points is given for Munich and Frankfurt a. M. or Munich and Leipzig.

⁶These include population, average age, gender, unemployment, wage, educational attainment, sector, firm size, GDP, rent and density. Evidence on the influence of the control variables from other research in the Appendix.

Table 2: Estimation results

	(1)	(2)	(3)
	Flow	Average flow	Flow
Distance	-0.712*** (0.0830)	-0.708*** (0.0835)	-0.659*** (0.0741)
Dialectsimilarity	0.691*** (0.0484)	0.689*** (0.0496)	0.684*** (0.0476)
Historical belonging	0.486*** (0.0385)	0.487*** (0.0391)	0.478*** (0.0381)
Constant	2.002*** (0.0804)	1.990*** (0.0807)	2.119*** (0.0730)
Region-time fixed effects	yes		yes
Region fixed effects		yes	
N	3689200	3689200	2699540
pseudo R^2	0.733	0.748	0.709

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

(1) Baseline PPML regression

(2) Estimated in means after Falck et al. (2012)

(3) Intensive margin of two-step method after Egger and Larch (2011)

The question of causality

Migration and place identity are intertwined and affect each other. In the following I try to solve this endogeneity problem.

I address the endogeneity concern by using the inner-German border as an instrument. For almost 30 years there was almost no migration that could shape place identity between East and West. Therefore, flows across the former border are much less driven by earlier migration. The disruption of the regional identity-migration cycle by the German wall hints to how the mechanism works.

Looking at the division also addresses another aspect that might drive movements of people: family or social ties. It is possible that historical belonging has no meaning for people's identity but has rather shaped the (family) distribution and now people move to where they have relatives or friends. However, during the existence of the German wall it was difficult to stay in contact with friends and family living on the other side of the wall. Of course there are numerous people who kept contact to their loved ones. But on average I expect movements between East and West after 1990 to be less strongly driven by earlier flows or relations because the connections were not as immediate.⁷

An estimation only regarding East-West and West-East migration flows shows a robust and significant effect of historical belonging, that is a bit less strong however (see

⁷Contrary to this, Burchardi and Hassan (2013) show that even after the division, social relations between East and West had a substantial economic effect. However, they do not compare their finding with the strength of social relations within the former East or West. I thus consider my assumption still reasonable.

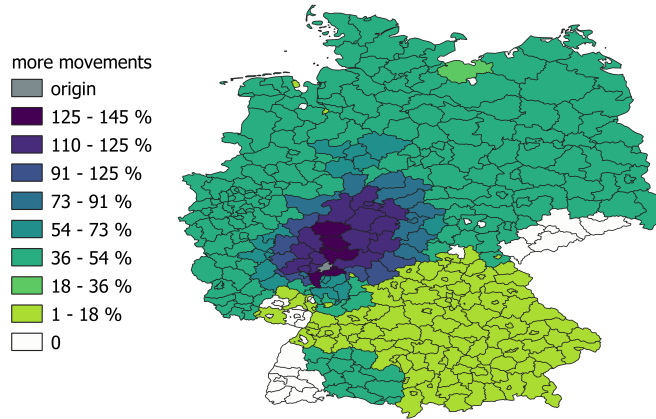


Figure 3: Magnitude of the effect of historical belonging by the example of Frankfurt

Table 6). Dialects have a lesser influence as well and are only significant to 5%.⁸ These results, displaying a small drop in the effect of historical belonging and a diminished effect of dialects, show that family bonds, social relations and/or earlier migration are likely to interfere in moving decisions. However, there is still a substantial effect of historical belonging. This points to a causal effect of identity on migration. Not only historic migration (disrupted during the separation of Germany) induces new migration by shaping place identity, but there also seems to be a more direct effect of historical connectedness.

In the following, several robustness tests are applied.

Sensitivity analysis

There are several aspects that potentially falsify my findings. I address them in the following subsections.

Specification problems

A relevant concern when regarding the influence of historical belonging and dialects simultaneously is a strong correlation between them. Really, the correlation is 0.5114 and therefore not strong enough to suspect serious multicollinearity.⁹

To see whether regional identity suffers from being a bad control variable I prove the explanatory power of distance, dialect similarity and historical belonging step by step. Outcomes are shown in Table 7. All three variables are highly significant and help to explain migration flows. As was to be expected, the introduction of the new measure of historical belonging lowers the influence of dialects because these aspects are intertwined. Nevertheless, both proxies are relevant and correlate by about 0.51 only. I therefore conclude that historical belonging serves as an additional explanatory factor and does not only shape migration through the channel of dialect similarity.

⁸When only including flows from 1995-2000, because later movements again might be driven by relations, the effect of historical belonging is robust (see Table 6).

⁹The correlation between distance and dialects is -0.63, between distance and historical belonging it is -0.48.

I furthermore perform a test whether the data is truncated or heteroscedastic, according to Silva and Tenreyro (2006). The outcome confirms the presence of heteroscedasticity; using a PPML is still adequate.

Another potential problem is non-linearity in the explanatory variables. Table 8 shows that regional identity affects migration more strongly when historical belonging has lasted longer. Distance has a negative effect for closer regions but is less prohibiting on longer distances (as in Schneider and Kubis, 2009). Dialect similarity likewise has a u-shaped effect. The explanatory power of the model goes up and the results generally confirm my base findings.

Variations of historical belonging

As an alternative measure for historical belonging, I use the time span length until most recently that regions uninterruptedly belonged together. This measure accounts for most recent belonging and how long it already lasts. By this it takes the possibility into account that migrants are not feeling connected to places that their region belonged to long ago. Results are found in Table 9 and are very similar to the base specification. The effect of historical belonging is slightly smaller, which shows that historical belonging also matters if it does not continue to exist. Additionally, I discount the base measure of historical belonging. In a linear manner, every year closer to the present counts one point more. With this, the effect of historical belonging is more strongly correlated with migration (see Table 10). The results of both these alternative specifications show that more current connections matter more but continuity is not as an important factor.

As another variation of the regional identity measure I regard Prussian provinces as individual political entities. Prussia was much larger than other states and had annexed a large part of Germany over time (see Figure 4). I reason that not all (annexed) regions identify with Prussia or other Prussian provinces. I therefore include the provinces separately as a reference category for regional identity. With this measure the estimation results are robust and show a slightly stronger effect of historical belonging and a slightly weaker impact of dialects (see Table 11). Dialects are more alike at the provincial level because throughout Prussia very different dialects were spoken.

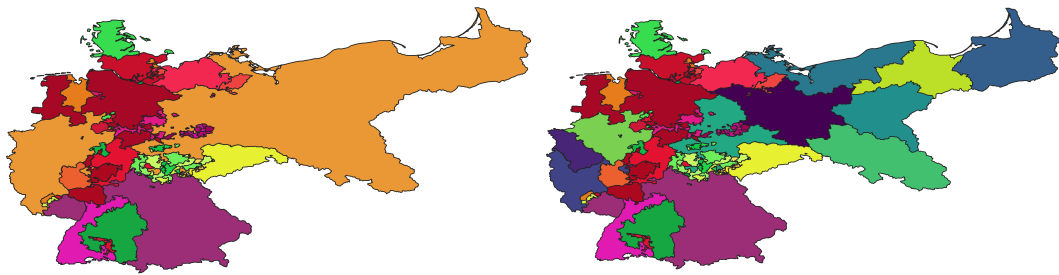


Figure 4: Prussia (1820) in orange (left) and divided in its provinces (right)
source: based on Kunz and Zipf (2008), MPIDR and CGG (2011a and 2011b)

Furthermore I add an East and West dummy to the base specification to account for the two former national states in Germany. Berlin is excluded in this specification due to its division. When including both dummies and only controlling for time fixed

effects¹⁰, the size of the historical belonging effect drops slightly and both dummies have a negative, insignificant effect (see Table 12). If including the East dummy only with time-region fixed effects I find a positive additional effect of moving within the East, significant to 5%. A separate West dummy does not have a significant effect. Because Berlin was divided after the second world war, assignment of the city to a specific territory is difficult. Also, the city has a special position as the capital. Dropping Berlin from the sample leads to similar results as in the base estimation (see Table 13). Distance has a slightly stronger influence and historical belonging is slightly weaker.

To prove that the mechanism of regional identity affecting migration has to do with feeling connected or not, in an alternative specification I refer to whether districts were opponents of war instead of sharing a history. Since 1820 the only war where German regions stood against each other was the German war in 1866. I derive whether today’s NUTS3 regions were opposing each other back then for every region pair. Including the resulting binary “war” variable in the estimation of migration flows shows that being wartime enemies 150 years ago still affects migration flows today (Table 14). The effect is less robust than historical belonging is but it only refers to one event.

Moves vs origin

The internal migration data set that I use represents all moves between 1995-2017. It thus does not allow tracking where people originally come from and how often individuals move around. Therefore, some of the observed moves are very likely the result of circle or return migration. Mühler and Opp (2004) expect that the earliest socialization forms individuals (and their regional identity) the strongest. In that case this aggregated data approach is biased by people moving from a place that is not their origin. Hence, I make use of a micro data set, that entails origin and current post codes of 57,000 individuals who participated in an internet personality test (Gosling and Potter, 2016). Because mostly young people and a higher share of women participated in this survey (Obschonka et al., 2019), findings are not necessarily representative for the whole population, however.

Participants are surveyed in the test for their post codes that are not always definitely matched with NUTS3 areas. For post codes that belong to more than one NUTS3 area I calculate the probability of stemming from the different NUTS3 areas according to the area sizes.

When only including individuals in the estimation whose current NUTS3 region does not match their NUTS3 region of origin (59% of the sample are dropped), estimation results show smaller effects of distance, dialects and historical belonging than in the base specification (see column (1) of Table 15). However, all variables are still significant to 1%.

Individuals are also observed in the data set if they do not migrate or have returned to their region of origin. Thus, the sample does not suffer from the selection bias problem. When including those who stayed within (or returned to) their NUTS3 region of origin, distance has an insignificant effect, whereas the effect of dialects is much stronger and that of historical belonging is a bit weaker compared to the base results (see column (2) of Table 15).

¹⁰Region fixed effects are omitted due to the specification.

Other explanations than regional identity

There are other factors that could explain my findings without referring to place identity stemming from historical belonging. Family or social ties have already been addressed as possible explanations.

Another concern regards federalism in Germany. Moving between federal states is administratively more difficult, for instance, moving children have to adjust to a new school syllabus. The measure of historical belonging also takes the federal states as regional entities into account because I expect a federal state (existing since 1949 in their majority) to shape place identity to some degree. Due to this however, the effect of the historical belonging variable potentially only reflects people’s unwillingness to move across federal state borders. I thus construct the measure of historical belonging without accounting for federal states (so counting only until 1949). Estimation results are very robust and historical belonging has a slightly weaker effect (see Table 16). This shows that the identity forming entities and connections are for a large part older than the federal states.

Additionally, I only regard movements within federal states. If the measure for historical belonging still shows an effect, then there is no administrative explanation for this. Results are found in Table 17 and show a robust effect of historical belonging and dialects, only distance has a stronger effect in this specification. However, when I look more closely at Baden-Wuerttemberg as an example for a federal state consisting of different historical entities, no significant effect of historical belonging is found. The effect of dialects is strong and seems to substitute for historical belonging (see Table 18). I hypothesize that in Baden-Wuerttemberg dialect closeness is more strongly determined by historical belonging and has a more important meaning for migration today.

A similar concern refers to the effect of historical belonging illustrating merely a preference for staying close to one’s place of birth. Close areas usually share a history. Therefore, in another robustness test I drop all moves within the same labour market region (Arbeitsmarktreion) or within a band of 50km or 100km from the sample. Results are found in table 19. They show that all effects drop in size but not in significance.

Likewise, city structure and the existence of (de)urbanisation has potentially biased the results. Acknowledging that major cities are usually differently perceived than their surroundings (“Munich is not Bavaria”), a further analysis drops moves to (and from) major cities (Table 20). I also control for people moving within the same spatial structure and for (de)urbanisation (Table 21). Results are robust to the base findings.

Regional identity and culture are naturally intertwined. In addition to dialect similarity I therefore control for other cultural characteristic in order to disentangle the effects: I regard the percentages of catholics and protestants in the regions and measure how much the two differ from one another. For this I compute the absolute difference between the regions for both religions and sum them up. The resulting measure shows a facet of cultural dissimilarity. In Table 22, it is seen that different religious denominations have a significant negative effect on migration. The impact of historical belonging however is still robust to the base findings. I similarly control for the difference in voting for the conservative party in Germany (CDU/CSU) between the NUTS3 regions as a measure of cultural dissimilarity. The effect is negative, as expected, but only significant to 10% (Table 23). Historical belonging as an explanatory variable is unaffected. Both analyses corroborate that regional identity is more than culture.

Conclusion

The concept of place identity states that people have an emotional attachment to places and integrate them in their identity. I show that this shapes the (economic) decisions whether and where to move. Regional identity is proxied by historical belonging, which is significantly positively connected to internal migration in Germany. Historical belonging measures whether today's districts belonged to the same (federal) state since 1820 and for how long this lasted. I connect the estimated effect to a regional identity, because the concept of "Heimat"/place identity and anecdotal evidence show that people refer to whether they feel at home in a region when they are moving. The correlation between internal migration and regional identity holds true if physical distance, dialect similarity and region specific characteristics are controlled for and with different estimation methods. Moreover, several robustness checks, like different specifications of the proxy of regional identity, confirm the outcome. Additionally, I regard other explanations than regional identity for how historical belonging affects today's migration: Both family ties and administrative hindrances can be disregarded as solely responsible for the estimation outcomes. In this paper I establish a correlation between migration and regional identity, that are expected to affect each other endogeneously. Estimations on the East-West and West-East flows after the division of Germany show, however, that historical ties also shape migration if earlier flows were interrupted for almost 30 years and did not further drive regional identity. Moving decisions, among other determinants, affect regional development. My findings therefore point to the relevance of place identity for the performance of regions.

Acknowledgements

I thank Nicola Sander and BBSR Wanderungsstatistik for providing the data on internal migration flows in Germany. Jens Südekum has kindly shared the dialect similarity index from Falck et al. (2012). Furthermore, Sam Gosling and Jeff Potter have helped by letting me use information from their survey data of the "The big five" internet personality test. I am grateful to Henri de Groot under whose supervision an early version of this research evolved and to Georg Hirte who supervises my PhD.

References

- Abraham, M. and Nisic, N. (2007). Regionale Bindung, räumliche Mobilität und Arbeitsmarkt-Analysen für die Schweiz und Deutschland. *Swiss Journal of Sociology*, 33(1).
- Adserà, A. (2015). Language and culture as drivers of migration. *IZA World of Labor*, 164.
- Adserà, A. and Pytliková, M. (2015). The Role of Language in Shaping International Migration. *Economic Journal*, 125(586):49–81.
- Anderson, J. E. (2011). The gravity model. *Annual Review of Economics*, 3(1):133–160.
- Anderson, J. E. and Van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review*, 93(1):170–192.
- Anton, C. E. and Lawrence, C. (2014). Home is where the heart is: The effect of place of residence on place attachment and community participation. *Journal of Environmental Psychology*, 40:451–461.
- Arntz, M. (2011). Mobilitätshemmnisse heterogener Arbeitskräfte in Deutschland. *Zeitschrift für Arbeitsmarktforschung*, 44(1-2):135–141.
- BBSR Wanderungsstatistik (2020). Wanderungsstatistik des Bundes und der Länder.
- BBSR Wohnungsmarktbeobachtung (2020). Angebotsmieten.
- Beaman, K. V. (2020). Identity and place in linguistic change across the life-span: The case of Swabian German (in press). In Ziegler, A., Edler, S., Kleczkowski, N., and Oberdorfer, G., editors, *Urban Matters. Current Approaches of International Sociolinguistic Research*. Benjamins, Amsterdam.
- Belot, M. and Ederveen, S. (2012). Cultural barriers in migration between OECD countries. *Journal of Population Economics*, 25(3):1077–1105.
- Bertelsmann-Stiftung (2008). Bürger und Föderalismus. Eine Umfrage zur Rolle der Bundesländer.
- Buenstorf, G., Geissler, M., and Krabel, S. (2016). Locations of labor market entry by German university graduates: is (regional) beauty in the eye of the beholder? *Review of Regional Research*.
- Burchardi, K. B. and Hassan, T. A. (2013). The economic impact of social ties: Evidence from German reunification. *Quarterly Journal of Economics*.
- Busch, O. and Weigert, B. (2010). Where have all the graduates gone? Internal cross-state migration of graduates in Germany 1984–2004. *The Annals of Regional Science*, 44(3):559–572.
- Cushing, B. and Poot, J. (2004). Crossing boundaries and borders: Regional science advances in migration modelling. *Papers in regional science*, 83(1):317–338.
- DaVanzo, J. (1980). *Micro economic approaches to studying migration decisions*. Rand, Santa Monica.
- Diercke Weltatlas (2009). Mitteleuropa - HRR 1648, Deutscher Bund 1815, Verwaltungsgliederung.
- Egger, P. and Larch, M. (2011). An assessment of the Europe agreements’ effects on bilateral trade, GDP, and welfare. *European Economic Review*, 55(2):263–279.
- Egger, P. H. and Tarlea, F. (2015). Multi-way clustering estimation of standard errors in gravity models. *Economics Letters*, 134:144–147.
- Falck, O., Heblich, S., Lameli, A., and Südekum, J. (2012). Dialects, cultural identity, and economic exchange. *Journal of Urban Economics*, 72(2-3):225–239.

- Falck, O., Lameli, A., and Ruhose, J. (2018). Cultural biases in migration: Estimating non-monetary migration costs. *Papers in Regional Science*, 97(2):411–438.
- Förtsch, M., Rösel, F., and Thum, M. (2019). Stärkung kommunaler Identität - Gutachten. Friedrich-Naumann-Stiftung für die Freiheit.
- Gosling, S. and Potter, J. (2016). Die ‘Big Five‘ Persönlichkeitstests. <http://de.outofservice.com/>.
- Hanns-Seidel-Stiftung (2009). Generationenstudie 2009: Heimatgefühl und Leben in Bayern.
- Helliwell, J. F. (1997). National borders, trade and migration. *Pacific Economic Review*, 2(3):165–185.
- Hernández Bernardo, B., Martín, A. M., Ruiz, C., and Hidalgo, M. d. C. (2010). The role of place identity and place attachment in breaking environmental protection laws. *Journal of Environmental Psychology*, 30(3):281–288.
- Hunt, J. (2006). Staunching emigration from East Germany: Age and the determinants of migration. *Journal of the European Economic Association*, 4(5):1014–1037.
- Kaasa, A., Vadi, M., and Varblane, U. (2016). A new dataset of cultural distances for European countries and regions. *Research in International Business and Finance*, 37:231–241.
- Kaplan, S., Grünwald, L., and Hirte, G. (2016). The effect of social networks and norms on the inter-regional migration intentions of knowledge-workers: The case of Saxony, Germany. *Cities*, 55:61–69.
- Krugman, P. (1991a). Increasing returns and economic geography. *Journal of Political Economy*, 99(3):483–499.
- Krugman, P. R. (1991b). *Geography and trade*. MIT press, Cambridge, MA.
- Kunz, A. and Zipf, A. (2008). Historical GIS Germany. <http://www.hgis-germany.de/>.
- Kupiszewski, M., Bucher, H., Durham, H., and Rees, P. (1998). Internal migration and regional population dynamics in Europe: German case study.
- Leibert, T. (2016). She leaves, he stays? Sex-selective migration in rural East Germany. *Journal of Rural Studies*, 43:267–279.
- Mayda, A. M. (2010). International migration: A panel data analysis of the determinants of bilateral flows. *Journal of Population Economics*, 23:1249–1274.
- McCallum, J. (1995). National borders matter: Canada-US regional trade patterns. *American Economic Review*, 85(3):615–623.
- Melzer, S. M. (2011). Reconsidering the effect of education on East–West migration in Germany. *European Sociological Review*, 29(2):210–228.
- Mitzscherlich, B. (1997). *"Heimat ist etwas, was ich mache". Eine psychologische Untersuchung zum individuellen Prozeß von Beheimatung*. Centaurus-Verl.-Ges., Pfaffenweiler.
- Molho, I. (2013). Theories of migration: A review. *Scottish Journal of Political Economy*, 60(5):526–556.
- MPIDR (Max Plank Institute for Demographic Research) and CGG (Chair for Geodesy and Geoinformatics University of Rostock) (2011a). MPIDR Population History GIS Collection (partly based on Bundesamt für Kartographie und Geodäsie 2011).
- MPIDR (Max Plank Institute for Demographic Research) and CGG (Chair for Geodesy and Geoinformatics University of Rostock) (2011b). MPIDR Population History GIS Collection (partly based on Hubatsch and Klein, 1975ff).
- Mühler, K. and Opp, K.-D. (2004). Die Ursachen regionaler und überregionaler Identität.

- fikation. In *Region und Nation*, pages 63–137. VS Verlag für Sozialwissenschaften, Wiesbaden.
- Newbold, K. B. (2019). The New Economic Geography in the Context of Migration. In Fischer, M. M. and Nijkamp, P., editors, *Handbook of Regional Science*. Springer, Berlin, Heidelberg.
- Obschonka, M., Wyrwich, M., Fritsch, M., Gosling, S. D., Rentfrow, P. J., and Potter, J. (2019). Von unterkühlten Norddeutschen, gemütlichen Süddeutschen und aufgeschlossenen Großstädtern: Regionale Persönlichkeitsunterschiede in Deutschland. *Psychologische Rundschau*.
- Oggenfuss, C. and Wolter, S. C. (2019). Are they coming back? The mobility of university graduates in switzerland. *Review of Regional Research*.
- Olivero, M. P. and Yotov, Y. V. (2012). Dynamic gravity: Endogenous country size and asset accumulation. *Canadian Journal of Economics*.
- Paasi, A. (2002). Bounded spaces in the mobile world: Deconstructing 'regional identity'. *Tijdschrift voor Economische en Sociale Geografie*, 93(2):137–148.
- Panzone, L., Di Vita, G., Borla, S., and D'Amico, M. (2016). When Consumers and Products Come From the Same Place: Preferences and WTP for Geographical Indication Differ Across Regional Identity Groups. *Journal of International Food and Agribusiness Marketing*, 28(3):286–313.
- Parikh, A., Van Leuvensteijn, M., and Others (2003). Internal migration in regions of Germany: A panel data analysis. *Applied Economics Quarterly*, 49(2):173–192.
- Peißker-Meyer, C. (2002). *Heimat auf Zeit*. transcript Verlag, Bielefeld.
- Peng, J., Strijker, D., and Wu, Q. (2020). Place Identity: How Far Have We Come in Exploring Its Meanings? *Frontiers in Psychology*, 11.
- Plaut, V. C., Markus, H. R., Treadway, J. R., and Fu, A. S. (2012). The Cultural Construction of Self and Well-Being: A Tale of Two Cities. *Personality and Social Psychology Bulletin*, 38(12):1644–1658.
- Proshansky, H. M., Fabian, A. K., and Kaminoff, R. (1983). Physical World Socialization of the Self. *Journal of Environmental Psychology*.
- Raagmaa, G. (2002). Regional identity in regional development and planning. *European Planning Studies*, 10(1):55–76.
- Ravenstein, E. G. (1885). The laws of migration. *Journal of the Statistical Society of London*, 48(2):167–235.
- Ríos, M. L. and Moreno-Jiménez, M. P. (2012). Place identity and residential satisfaction: Differences between native and immigrant populations. *Psychology*, 3(1):75–86.
- Sander, N. (2014). Internal migration in Germany, 1995-2010: New insights into east-west migration and re-urbanisation. *Comparative Population Studies*, 39(2).
- Schneider, L. and Kubis, A. (2009). Are there gender-specific preferences for location factors? A Grouped Conditional Logit-Model of interregional migration flows in Germany. Technical report, IWH Discussion Papers.
- Sedlacek, S., Kurka, B., and Maier, G. (2009). Regional identity: a key to overcome structural weaknesses in peripheral rural regions? *European Countryside*, 1(4):180–201.
- Shepherd, B. (2016). *The Gravity Model of International Trade: A User Guide - Chapter 4: Alternative Gravity Model Estimators*. United Nations ESCAP, <https://www.unescap.org/sites/default/files/6%20->

- 204.AlternativeGravityModelEstimators_0.pdf (accessed 10 June 2019).
- Silva, J. M. C. S. and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 88(4):641–658.
- Stawarz, N. and Sander, N. (2020). The Impact of Internal Migration on the Spatial Distribution of Population in Germany over the Period 1991-2017. *Comparative Population Studies*, 44.
- Sturm, R. (2016). Regionale politische Kulturen im deutschen Föderalismus. In Werz, N. and Koschkar, M., editors, *Regionale politische Kultur in Deutschland*. Springer Fachmedien, Wiesbaden.
- Wehling, H.-G. (2013). *Die deutschen Länder: Geschichte, Politik, Wirtschaft*. Springer-Verlag.

Appendix

Control variables for internal migration

In the following the control variables of the model are reasoned. They are used as substitutes for region-time fixed effects in a robustness test. Data sources are found in Table 3.

Table 3: Data sources

data	source	spatial unit	years
migration flows	BBSR Wanderungsstatistik	NUTS 3	1991-2017
distance	GIS (own calculation)	NUTS 3	constant
culture	dialect similarity index (Falek et al., 2012)	NUTS 3	constant
wages	BBSR Bonn (Inkar)	labour market region	2000-2016
income	BBSR Bonn (Inkar)	labour market region	2000-2016
unemployment	BBSR Bonn (Inkar)	labour market region	1998-2017
population	Volkswirtschaftliche Gesamtrechnung der Länder (VGRdL)	NUTS 3	1995-2017
average age	BBSR Bonn (Inkar)	NUTS 3	1995-2017
gender	BBSR Bonn (Inkar)	NUTS 3	1995-2017
educational attainment of employees	labour market region	NUTS 3	2012-2017
GDP pc	BBSR Bonn (Inkar)	labour market region	2000-2017
sectoral composition	BBSR Bonn (Inkar)	labour market region	2008-2017
firm size	BBSR Bonn (Inkar)	labour market region	2006-2015
density	BBSR Bonn (Inkar)	NUTS3	1997-2017
rents	BBSR Wohnungsmarktbeobachtung	NUTS 3	2004-2019

The focus of analysis in this work lies on interregional migration between NUTS3 regions which are equivalent to the German districts (“Kreise”). Data on control variables is available for different time periods such that analysis is only possible for 2012-2015.

The explanatory control variables, used instead of region-time fixed effects, are as follows:

From the disequilibrium theory a positive influence of income difference between destination and origin is predicted. Likewise the difference in unemployment should affect flows negatively (compare for both Arntz, 2011; Kupiszewski et al., 1998; Parikh et al., 2003; Schneider and Kubis, 2009). Because expectedly not only the labour market condition of the districts matter but due to commuting possibilities rather that of the “Arbeitsmarktregion” (labour market region), I control for wages and unemployment at this level. Other labour market characteristics that might affect migration are also controlled for at the same level: Educational attainment of employees (expected to positively influence migration levels; DaVanzo, 1980), firm size, sectoral composition and GDP per capita.

Due to the relevance of individual characteristics in migration decisions, aggregate data on age and binary gender is also employed (according to Busch and Weigert, 2010;

Leibert, 2016; Melzer, 2011; Hunt, 2006). Using data on an individual level would be more adequate but averages work as proxies. Education data is not available at a spatially refined level.

Furthermore have density (see for urban-rural dynamics e.g. Sander, 2014) and rents an expected effect on migration.

Estimation tables

Table 4: Estimation with control variables instead of fixed effects

	(1) Flow
Distance	-0.729*** (0.0851)
Dialect similarity	0.692*** (0.0412)
Historical Belonging	0.483*** (0.0369)
Population (orig)	-0.000870 (0.0309)
Average age (orig)	-0.164*** (0.0300)
Share females (orig)	-0.0124 (0.00977)
Unemployment (orig)	0.114*** (0.0298)
Wage (orig)	-0.00105 (0.0173)
Rent (orig)	0.0439*** (0.0126)
Density (orig)	-0.167* (0.0954)
Share unskilled (orig)	-0.0229** (0.00981)

Share professionals (orig)	-0.0276** (0.0138)
Share academics (orig)	0.0685*** (0.0237)
Primary sector (orig)	0.0180 (0.0267)
Secondary sector (orig)	0.0691 (0.183)
Tertiary sector (orig)	0.108 (0.180)
Small firms (orig)	0.00675 (0.00886)
Medium firms (orig)	-0.000184 (0.00749)
Large firms (orig)	0.0104** (0.00527)
GDP (orig)	0.0143 (0.0124)
Unemployment (dest)	0.00554 (0.0405)
Wage (dest)	-0.00391 (0.0228)
Rent (dest)	-0.0543*** (0.0159)
Density (dest)	-0.110 (0.127)
Share unskilled (dest)	-0.0154 (0.0138)
Share professionals (dest)	0.0263 (0.0176)
Share academics (dest)	-0.0918*** (0.0288)
Primary sector (dest)	-0.000732 (0.0406)

Secondary sector (dest)	-0.118 (0.311)
Tertiary sector (dest)	-0.130 (0.308)
Small firms (dest)	-0.00805 (0.00866)
Medium firms (dest)	0.0104 (0.00984)
Large firms (dest)	0.00355 (0.00832)
GDP (dest)	-0.0215 (0.0195)
Constant	2.293*** (0.142)
Region-time fixed effects	no
N	638400
pseudo R^2	0.756

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

Observation number limited due to data availability

Table 5: Two-step estimation, extensive margin: 2000 and 2010 as examples

	(1)	(2)
	2000	2010
Distance	-0.392*** (0.00844)	-0.451*** (0.00857)
Dialect similarity	0.198*** (0.00812)	0.153*** (0.00824)
Historical belonging	0.169*** (0.00612)	0.178*** (0.00608)
Constant	2.153*** (0.222)	5.614*** (0.457)
Region fixed effects	yes	yes
N	158403	158802
pseudo R^2	0.301	0.294

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

The number of observations differs across years because fixed region effects predict some of the outcomes perfectly, and the number of pairs for which this happens changes over the years. P-value of LR-tests is 0.0000.

Table 6: Causality and family ties: Pooled East-West and West-East flows

	(1)	(2)
	Flow	Flow
Distance	-0.896*** (0.0684)	-1.013*** (0.0850)
Dialect similarity	0.0990** (0.0463)	0.0647 (0.0575)
Historical belonging	0.362*** (0.0678)	0.362*** (0.0662)
Constant	2.268*** (0.0118)	2.326*** (0.0128)
Region-time fixed effects	yes	yes
N	1132704	295488
pseudo R^2	0.560	0.536

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

(1) Are all flows across the former border

(2) Are only flows between 1995-2000

Table 7: Distance, dialect similarity and historical belonging separately

	(1)	(2)	(3)	(4)	(5)
	Flow	Flow	Flow	Flow	Flow
Distance	-1.959*** (0.121)	-0.999*** (0.103)	-0.712*** (0.0830)	-1.327*** (0.115)	
Dialect similarity		0.833*** (0.0508)	0.691*** (0.0484)		
Historical Belonging			0.486*** (0.0385)	0.634*** (0.0446)	1.434*** (0.0431)
Constant	1.915*** (0.146)	2.004*** (0.0968)	2.002*** (0.0804)	1.983*** (0.108)	2.346*** (0.0543)
Region-time fixed effects	yes	yes	yes	yes	yes
N	3689200	3689200	3689200	3689200	3689200
pseudo R^2	0.665	0.713	0.733	0.701	0.600

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

Table 8: Non-linearities in explanatory variables

	(1) Flow
Distance	-6.987*** (0.184)
Distance_2	13.87*** (0.505)
Distance_3	-7.834*** (0.377)
Dialect similarity	-0.565*** (0.0826)
Dialect similarity_2	0.659*** (0.0761)
Historical belonging	-0.166*** (0.0580)
Historical belonging_2	0.292*** (0.0555)
Constant	2.191*** (0.0363)
Region-time fixed effects	yes
N	3689200
pseudo R^2	0.851

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

Table 9: Alternative measure: Length of most recent belonging

	(1)
	Flow
Distance	-0.730*** (0.0769)
Dialect similarity	0.686*** (0.0475)
Historical belonging (continuous until present)	0.367*** (0.0216)
Constant	2.076*** (0.0696)
Region-time fixed effects	yes
N	3689200
pseudo R^2	0.742

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

Table 10: Alternative measure: Discounted historical belonging

	(1)
	Flow
Distance	-0.642*** (0.0840)
Dialect similarity	0.682*** (0.0487)
Historical belonging (discounted)	0.522*** (0.0417)
Constant	2.016*** (0.0748)
Region-time fixed effects	yes
N	3689200
pseudo R^2	0.738

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

Table 11: Alternative measure: Prussia's provinces included as individual units

	(1) Flow
Distance	-0.601*** (0.0958)
Dialect similarity	0.657*** (0.0486)
Historical belonging (Provinces)	0.493*** (0.0395)
Constant	2.097*** (0.0771)
Region-time fixed effects	yes
N	3689200
pseudo R^2	0.740

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

Table 12: Adding an East and West dummy

	(1) Flow
Distance	-1.126*** (0.111)
Dialect similarity	0.484*** (0.0535)
Historical Belonging	0.337*** (0.0488)
East	-0.237 (0.276)
West	-0.406 (0.291)
Constant	1.614*** (0.313)
Time fixed effects	yes
N	3689200
pseudo R^2	0.447

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

Region fixed effect are dropped due to the specification

Table 13: Dropping Berlin from the sample

	(1)
	Flow
Distance	-0.811*** (0.0662)
Dialect similarity	0.686*** (0.0488)
Historical belonging	0.461*** (0.0368)
Constant	1.746*** (0.0718)
Region-time fixed effects	yes
N	3670800
pseudo R^2	0.726

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

Table 14: Alternative measure: Opponents in the German war

	(1)
	Flow
Distance	-0.971*** (0.101)
Dialect similarity	0.818*** (0.0515)
War	-0.277*** (0.0594)
Constant	2.107*** (0.0943)
Region-time fixed effects	yes
N	3689200
pseudo R^2	0.715

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

War is a binary variable (1 for war opponents)

Table 15: Estimation of micro data set entailing origin and current place of individuals

	(1)	(2)
	Micro flow	Micro flow
Distance	-0.503*** (0.0899)	-0.0832 (0.169)
Dialect similarity	0.534*** (0.0345)	2.042*** (0.0750)
Historical belonging	0.432*** (0.0313)	0.287*** (0.0539)
_cons	-1.560*** (0.0700)	-2.977*** (0.267)
Region-time fixed effects	yes	yes
N	160400	160801
pseudo R^2	0.504	0.719

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

(1) Are only people who live in a different place

(2) Includes those that stayed or returned

Table 16: Alternative measure: Historical belonging until 1949

	(1)
	Flow
Distance	-0.849*** (0.0899)
Dialect similarity	0.735*** (0.0483)
Historical belonging (until 1949)	0.374*** (0.0332)
Constant	1.989*** (0.0879)
Region-time fixed effects	yes
N	3689200
pseudo R^2	0.724

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

Table 17: Flows within federal states only

	(1) Flow
Distance	-4.362*** (0.186)
Dialect similarity	0.491*** (0.0444)
Historical belonging	0.277*** (0.0605)
Constant	-3.401*** (0.277)
Region-time fixed effects	yes
N	439392
pseudo R^2	0.842

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

Table 18: Flows within Baden-Wuerttemberg (as example)

	(1) Flow	(2) Flow	(3) Flow
Distance	-3.474*** (0.276)	-3.469*** (0.271)	-4.685*** (0.363)
Dialect similarity	0.764*** (0.0653)	0.752*** (0.0601)	
Historical belonging	-0.0236 (0.0802)		0.269*** (0.0735)
Constant	-1.780*** (0.472)	-1.790*** (0.474)	-2.803*** (0.579)
Region-time fixed effects	yes	yes	yes
N	43516	43516	43516
pseudo R^2	0.822	0.822	0.794

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

Table 19: Flows dropped within close regions

	(1)	(2)	(3)
	Flow	Flow	Flow
Distance	-0.654*** (0.0640)	-0.583*** (0.0464)	-0.446*** (0.0384)
Dialect similarity	0.535*** (0.0415)	0.220*** (0.0234)	0.107*** (0.0171)
Historical belonging	0.435*** (0.0312)	0.312*** (0.0208)	0.224*** (0.0171)
Constant	2.204*** (0.0566)	2.597*** (0.0199)	2.666*** (0.00582)
Region-time fixed effects	yes	yes	yes
N	3677792	3603410	3378930
pseudo R^2	0.720	0.725	0.689

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

(1) Flows dropped within labour market regions

(1) Flows dropped within 50km

(1) Flows dropped within 100km

Table 20: Controlling for major cities

	(1)	(2)
	Flow	Flow
Distance	-0.904*** (0.0668)	-0.973*** (0.0694)
Dialect similarity	0.671*** (0.0573)	0.565*** (0.0532)
Historical belonging	0.485*** (0.0404)	0.472*** (0.0338)
Constant	1.485*** (0.0623)	1.155*** (0.0623)
Region-time fixed effects	yes	yes
N	3072800	2558106
pseudo R^2	0.716	0.665

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

(1) Flows directed from major cities are dropped

(1) Flows directed to and from major cities are dropped

Table 21: Controlling for the same spatial structure and for (de)urbanisation

	(1) Flow	(2) Flow
Distance	-0.708*** (0.0825)	-0.702*** (0.0789)
Dialect similarity	0.692*** (0.0488)	0.692*** (0.0486)
Historical belonging	0.487*** (0.0387)	0.485*** (0.0380)
same Kreistyp	0.0760** (0.0306)	
Urbanisation		-0.398*** (0.0846)
Deurbanisation		0 (.)
Constant	1.980*** (0.0773)	2.086*** (0.0861)
Region-time fixed effects	yes	yes
N	3689200	3689200
pseudo R^2	0.734	0.735

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

(1) Controls for same spatial structure in origin and destination according to the BBSR classification

(2) Controls for (de)urbanisation moves

Table 22: Controlling for another cultural factor: Religion

	(1) Flow
Distance	-0.616*** (0.0861)
Dialect similarity	0.683*** (0.0474)
Historical belonging	0.476*** (0.0383)
Religion	-0.156*** (0.0183)
Constant	1.965*** (0.0742)
Region-time fixed effects	yes
N	3689200
pseudo R^2	0.738

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

Religion measures the absolute, summed difference
between the shares of catholics and protestants
for each two regions

Table 23: Controlling for another cultural factor: Conservativeness

	(1)
	Flow
Distance	-0.709*** (0.0836)
Dialect similarity	0.691*** (0.0481)
Historical belonging	0.484*** (0.0393)
Conservative	-0.0566* (0.0319)
Constant	1.977*** (0.0819)
Region-time fixed effects	yes
N	962400
pseudo R^2	0.737

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

All explanatory variables are standardized

Conservative measures the difference in voting for the conservative parties for every region pair