Dialects, Cultural Identity, and Economic Exchange

Oliver Falck[+], Stephan Heblich[*], Alfred Lameli[‡], Jens Südekum[†]

October 2010

[+] Ifo Institute for Economic Research, Poschingerstr. 5, D-81679 Munich (Germany), Phone: +49 89 9224 1370, Fax: +49 89 9224 1460, Email: falck@ifo.de, CESifo and Max Planck Institute of Economics.

[*] Max Planck Institute of Economics, Entrepreneurship, Growth, and Public Policy Group, Kahlaischestr. 10, D-07745 Jena (Germany), Phone: +49 3641 686 733, Fax: +49 3641 686 710, Email: heblich@econ.mpg.de.

[‡] Research Centre *Deutscher Sprachatlas*, Hermann-Jacobsohn-Weg 3, D-35032 Marburg (Germany), Phone: +49 6421 28 22 482, Fax: +49 6421 28 28 936, Email: lameli@staff.uni-marburg.de.

[†] Mercator School of Management, University of Duisburg-Essen, Lotharstraße 65, D-47057 Duisburg (Germany), Phone: +49 203 379 2357, Fax: +49 203 379-1786, Email: jens.suedekum@uni-due.de

Dialects, Cultural Identity, and Economic Exchange

Abstract

It has long been argued that economic phenomena are affected by culture. However, the causal effect of cultural ties on economic exchange is difficult to identify, chiefly because cultural ties are endogenous to the current level of economic exchange, and because it is hard to separate *culture* from other influences. In this paper, we address these issues by using a novel measure for cultural identity—historical dialect differences across regions of the same country. We evaluate linguistic micro-data from a unique language survey conducted between 1879 and 1888 in about 45,000 German schools. The recorded geography of dialects comprehensively portrays local cultural differences that have been evolving for centuries and provides an ideal opportunity to isolate cultural costs from other barriers to economic exchange. In a gravity analysis, we then show that cross-regional migration flows in the period 2000–2006 are positively affected by historical dialect similarity, a finding that indicates highly time-persistent cultural borders that impede economic exchange even at a fine geographical scale.

1. Introduction

Scholars from various disciplines have long argued that economic phenomena are affected by culture, but quantitative research on the causal effects of cultural ties on economic exchange has started only very recently. *Culture* is not an easily operational concept, however, and is thus difficult to measure. Specifically, proxies for cultural ties: i) are often available only at a high level of aggregation, typically for different countries; ii) may also capture other effects on economic outcomes, such as institutional or information barriers; iii) do not provide a comprehensive picture of cultural identity, but capture only single aspects such as religion or ethnicity; and iv) are often endogenous to the level of economic exchange.

In order to address those problems, we develop a novel approach for studying the economic implications of cultural ties. Our units of analysis are small-scale regions of the same country (Germany) across which institutional or information barriers are much lower than between countries. Based on linguistic micro-data from the late 19th century, we explore the pattern of dialects across those regions and show that these language structures give a unique and comprehensive portray of persistent cultural ties. We then investigate how the historical dialect similarity affects current cross-regional economic exchange, specifically, migration flows. Our central finding is that contemporaneous migration is significantly positively affected by the similarity of the dialects prevalent in the source and the destination area more than 120 years ago, even after controlling for geographical distances and regional fixed effects. Decreasing geographical distance by 100 kilometers raises gross regional migration flows by roughly 41% directly, and by another 2-3% indirectly through increased dialect similarity. Various robustness checks confirm that this latter effect can be interpreted as a causal impact of cultural ties on regional migration in Germany. In other words, our results

indicate that internal migration in Germany is impeded by intangible and highly time-persistent cultural borders, even at this fine geographical scale.

From a linguistic point of view, dialects are local variants of the same language. They can vary substantially in terms of pronunciation, vocabulary, or grammar, but despite their peculiarities, dialects are not major barriers to communication in the same way that completely different languages are (Crystal, 1987).[1] We evaluate micro-data derived from an encompassing language survey conducted by the linguist Georg Wenker between 1879 and 1888. This survey was intended to provide a detailed inquiry into language variation within the newly created German Empire, which was comprised of a large number of formerly autonomous princedoms and kingdoms. The survey includes detailed phonological and grammatical attributes of the languages spoken by pupils in about 45,000 schools across the German Empire.[2] Using these data, we construct a dialect similarity matrix for the current 439 German districts (NUTS3 regions), where the characterization of each district's dialect is based on 383 linguistic features having to do with the pronunciation of consonants and vowels as well as with grammar.

Below we provide several examples of our data showing that the geography of dialects is the outcome of a century-long evolutionary process and—almost like a genome—stores information about historical interactions across the German regions that would be very difficult, if not impossible, to capture without linguistic data. That such data are useful to comprehensively portray differences across populations is not a new idea; in fact, already Charles Darwin pointed out that language structures are extremely powerful in this respect when he wrote, in his seminal book, *Origin of Species*:

---

[1] For example, the dialect spoken in Liverpool ("Scouse") is quite distinct from the English spoken in most parts of the United Kingdom, including the neighboring regions of Chesire and Lancashire. Nevertheless, "Scousers" can still easily engage in conversation with people from other areas of the country.

[2] To this day, the Wenker survey is the most complete documentation ever of a nation's language and has defined standards in the linguistics discipline (for a detailed introduction, see Lameli 2008). Similar data are not available for other countries or languages at a comparable degree of detail.

If we possessed a perfect pedigree of mankind, a genealogical arrangement of the races of man would afford the best classification of the languages now spoken around the world; and if all extinct languages, and all intermediate and slowly changing dialect, were to be included, such an arrangement would be the only possible one. (cited after Cavalli-Sforza 2000:167).

The data from the Wenker survey, therefore, seem to provide a unique opportunity to measure deep cultural differences at an unusual degree of detail, the region-pair level. Based on a gravity model, we then investigate to what extent the historical dialect (cultural) differences affect regional gross migration flows during the period 2000–2006. Such contemporaneous migration flows are a well suited outcome variable, because migration (also across regions) typically occurs only rarely in an individual's lifetime. Cultural factors are thus likely to influence such decisions even more strongly than, say, they would the decision to trade goods or to conduct financial transactions with someone from a different region.

Research on internal migration consistently shows that local economic conditions, such as wages, unemployment, housing prices, etc., are key drivers of individual location decisions. In our gravity model for migration, we control for contemporaneous differences in economic prosperity with source and destination area fixed effects. These fixed effects also capture other (contemporaneous and historical) unobservable region-specific attributes, such as amenities, institutions, or historical factors that vary only at the regional level. It is also frequently noted in the migration literature that gross flows are two-directional. That is, the direction of migration is not always from poor to rich places, but also the other way around (Schwartz 1973; Greenwood 1975). Moreover, gross flows are often relatively small, especially in European countries where regional labor mobility is lower than it is in the United States (Hunt 2006). These observations suggest that migration is accompanied by significant costs other than and in addition to the strictly economic ones. In a classic article, Sjaastad (1962) uses the term "psychic costs of migration," which include the costs of leaving "familiar sur-

roundings" and therefore encapsulate cultural differences across regions that affect mobility decisions. Our region-pair-specific dialect similarity matrix allows us, as one of the first attempts in the literature, to directly proxy the "psychic mobility costs" and to quantify their effect on regional migration flows.

Our study is related to a recent line of research on the economic effects of cultural differences at the international level. Guiso *et al.* (2009) show that trade and investment flows across countries are impeded by intangible borders that stem from a lack of bilateral trust (measured with data from the *World Value Survey*). Bilateral trust is explained, inter alia, by cultural differences across countries, which are in turn measured by differences in religious, genetic, and somatic traits. Our study adds to this literature by showing that cultural borders to economic exchange also exist *within* nations and, hence, on a much finer geographical scale.[3] Furthermore, our study allows us to abstract from distortions such as institutions or information barriers, which are difficult to separate from the effect of culture, as those other distortions surely matter more at the international than at the regional level.

Tabellini (2010) studies the interaction of culture and institutions in determining output differences across European regions. Recognizing that contemporaneous cultural values are endogenous to current economic outcomes, he instruments culture with historical regional variables on literacy rates and political institutions. In our study, we directly observe past cultural differences across regions by relying on dialects measured in the 19th century and, therefore, face no comparable simultaneity problem. For a subset of German regions we can even test if the historical dialect differences are a relevant measure for contemporaneous dialect (cultural) differences. Using linguistic data from a recent language survey for 69 Bavarian regions, we illustrate that the relative linguistic differences across regions that Georg

---

[3] Cultural and genetic differences are also put in perspective by Desmet *et al.* (2008), who show that countries with more distant gene profiles exhibit stronger differences in cultural values. Spolaore and Wacziarg (2009) find a positive relationship of genetic and current income differences across countries. These papers thus emphasize that genetic similarity stimulates economic exchange via a cultural channel. We obtain a consistent result for linguistic similarity on a much finer spatial scale.

Wenker recorded in the 19th century have remained almost unchanged until today, i.e., that dialect differences are highly time-persistent. We then use the historical dialects as an instrument for current dialect differences in a two-stage specification, in order to establish the causality of the effect of cultural ties on current regional migration.

We also take into account two other possible sources of endogeneity. First, we address omitted variable bias and show that dialect similarity is not confounded with other types of region-pair-specific congruencies, like a common religious or political history, a similar industry structure, or connectedness through ancient trading routes. Consistent with the argument by Charles Darwin, our study indeed suggest that linguistic data capture an entirety of historical imprints and otherwise immeasurable aspects of cultural identity. Second, we address the possibility of a spurious correlation that could result from a persistence of migration flows. If this was the case, linguisitic similarity could be the outcome of previous migration flows, rather than a causal influence on current migration. Exploiting the quasi-natural experiment of German Reunification, we analyze if network effects may have led to persistence in migration flows and, in turn, to dialect assimilation. Furthermore, by focusing on subsamples of regions with similar geological features, we address the possibility that migration flows are driven by persistent omitted factors over the very long run.

Finally, our paper is more broadly related to the extensive literature about language and economics which has mostly discussed language acquisition by foreigners or the coexistence of domestic and foreign languages within one country.[4] Our focus in this paper is different in that we study regional variation of the *same* language. Our findings are thus unlikely to be due to a similar transaction cost mechanism as in Lazear's (1999) model, where individuals can conduct economic transactions only when they speak a common language. Dialect dif-

---

[4] Important contributions to the literature on language acquisition by immigrants and multilingual countries include Chiswick (1991), Rauch (1999), Rauch and Trindade (2002), Alesina and La Ferrara (2005), and Melitz (2008). Ginsburgh and Weber (2010) comprehensively review the literature on linguistic diversity.

ferences matter for internal migration decisions not because people would be literally unable to communicate in other regions, but because they have a preference for living in a culturally familiar environment.[5]

The remainder of this paper is organized as follows. In Section 2 we describe our linguistic data and discuss the meaning of dialects, especially in the historical context of our study. Section 3 sets out a simple gravity model for current migration flows that serves as the underlying framework for the empirical analysis. Section 4 presents our baseline results, and Section 5 is devoted to various robustness checks. Section 6 concludes.

2. Background and data

2.1. *Historical background and the measurement of linguistic characteristics*

In the centuries following Charlemagne, France, Spain, England, and Habsburg Austria developed into states where power was wielded by a centralized sovereign. In contrast, the Holy Roman Empire became increasingly fragmented. When the Treaty of Westphalia ended the Holy Roman Empire in 1648, what we know as Germany today was comprised of hundreds of sovereign kingdoms, principalities, and dukedoms. This fragmentation continued until the German Empire (*Deutsches Reich*) was established in the late 19th century.

Georg Wenker conducted his language survey shortly after the Empire was established. Between 1879 and 1888, he asked teachers and pupils in more than 45,000 schools to translate 40 German sentences into their local dialect. These sentences were especially designed to reveal specific dialect characteristics. The survey covered the entire area of the Empire and revealed pronounced differentiation of local languages, since at that time (more so than today) dialects were the people's common everyday speech.

[5] The only other economic study we are aware of that exploits linguistic variation within the same language is Grogger (2010). He finds that black workers in the United States systematically earn lower wages if their speech can be distinctively identified as African-American English. Our focus is different in that we are not interested in labor market discrimination against particular dialects (or sociolects), but in the effects of cultural similarity—as measured by dialect similarity—on economic exchange.

Wenker's surviving material contains millions of phonological and grammatical observations in the form of handwritten protocols of the language characteristics recorded in the individual schools (see Figure 1a for an example). These raw data were integrated by Wenker and collaborators into a linguistic atlas of the German Empire (*Sprachatlas des Deutschen Reichs*). The *Sprachatlas* was developed between 1889 and 1923 and contains more than 1,600 hand-drawn maps showing the detailed geographical distribution of particular language characteristics across the German Empire (see Figure 1b for an example). In an evaluation process that spanned several decades, Ferdinand Wrede, Wenker's successor in the linguistic atlas project, determined 66 prototypical characteristics that are most relevant for the structuring of the German language area.[6] Each of those characteristics has to do with the pronunciation of consonants and vowels as well as with grammar. An individual map exists for each of those 66 linguistic attributes.[7]

[FIGURES 1a AND 1b HERE]

We matched these 66 thematic maps with Germany's current administrative classification scheme. The Federal Republic of Germany currently consists of 439 districts (*Landkreise*); however, the linguistic maps from the *Sprachatlas* do not conform to this classification system. We therefore use GIS (Geographical Information System) technology to juxtapose digitized versions of these linguistic maps and the map of the current administrative districts. We then quantify the dialect of each district in the form of binary variables.

The following example illustrates this approach. One of the linguistic attributes is the German word for *pound*. Depending on the dialect, it is pronounced as "Pfund," "Pund," or "Fund." The corresponding map in the *Sprachatlas* shows the variant "Fund" mostly in the eastern parts of Germany, "Pund" mostly in the northern areas, and "Pfund" mostly in the

---

[6] Wrede combined local extractions of variants to a dialect classification (see Wrede *et al.* 1927–1956, map 56). One advantage of this classification over more recent categorizations of the Wenker data (e.g., Wiesinger 1983b) is that it lends itself quite easily to a mathematical representation of dialects (see below).

[7] All hand-drawn maps are published online as the 'Digitaler Wenker-Atlas' (DiWA), see http://www.diwa.info.

southern parts. These variants are then transferred into a binary coding of the type: "Fund" = {1 0 0}; "Pund" = {0 1 0}; "Pfund' = {0 0 1}. Comparing the individual linguistic map for the word *pound* and the current administrative map of Germany, we assign one of these codes to each of the 439 districts. This approach is unambiguous when there is no intra-regional variation of this particular language characteristic, i.e., when the entire area of some district r exhibited the same pronunciation according to the map in the *Sprachatlas*. Typically this has been the case. However, the spatial distribution of this particular language attribute and the current boundaries of the districts are not in all cases perfectly coincident. If we found intra-regional variation of pronunciation, we then chose the most frequent variant within the district as representative. The entire matching procedure was accompanied by several linguistic plausibility tests and cross-checks with the underlying raw data on the phonetic protocols from the Wenker survey.

Repeating this procedure for all 66 language characteristics, which can have between 2 and 18 realizations, we end up with K=383 binary variables representing the dialect that was spoken in the area of a district in the late 19th century. More formally, the historical dialect of the current district r is represented by a vector $\mathbf{i}^r = \left\{ i_1^r, i_2^r, \cdots, i_K^r \right\}$ of length K=383, where each vector element is a binary variable [0,1]. Using these data, we can then construct a dialect similarity matrix across all R=439 regions as follows: consider any two German regions r and s whose historical dialects are represented by $\mathbf{i}^r = \left\{ i_1^r, i_2^r, \cdots, i_K^r \right\}$ and $\mathbf{i}^s = \left\{ i_1^s, i_2^s, \cdots, i_K^s \right\}$, respectively. To quantify the overlap of these two vectors, we use a simple count similarity measure, namely $\ell_{rs} = \mathbf{i}^r \times \mathbf{i}^s$, where $0 \leq \ell_{rs} \leq 66$ for $r \neq s$ and $\ell_{rr} = 66$. The resulting matrix across all regions then has dimension 439×439 with generic elements ℓ_{rs}. As a robustness check we also considered two different similarity indices. First, Jaccard's (1901) index is computed as follows: Given the two vectors $\mathbf{i}^r$ and $\mathbf{i}^s$, let M_{11} be the

number of vector columns where both $\mathbf{i}^r$ and $\mathbf{i}^s$ have the value 1, M_{10} the number of cases where $\mathbf{i}^r$ has a 1 and $\mathbf{i}^s$ has a 0, M_{01} the number of cases where $\mathbf{i}^r$ has a 0 and $\mathbf{i}^s$ has a 1, and M_{00} the number of cases where both vectors have a 0. The Jaccard index is then defined as $M_{11}/(M_{11}+M_{10}+M_{01})$. Second, Kulczynski's (1927) similarity index is defined as $\frac{1}{2}\cdot\left[M_{11}/(M_{11}+M_{10})+M_{11}/(M_{11}+M_{01})\right]$. Both indices range between zero and one, with a higher value indicating a higher degree of dialect similarity. Table 1 provides some summary statistics. For our preferred count similarity measure, the average number of linguistic correspondences across all regional pairs is $\overline{\ell}_{rs}=32.6$ with a standard deviation equal to 10.53.

[Table 1 HERE]

2.2. *What does dialect similarity capture?*

Nations are not monolithic linguistically. Typically, there are hundreds of regional dialects within the same language. Such phonological and grammatical variations across space are by no means random, but they have developed over centuries in a process of linguistic evolution and, therefore, reflect various imprints from the past. Charles Darwin was already aware of this point, and a vast subsequent literature has further dwelled on the fact that both genes and languages are correlated and the product of evolution.[8] As such, both exhibit an enormous degree of persistence. Yet, anthropologists such as Cavalli-Sforza (2000) argue that language differences across groups tend to be better visible than genetic ones, as they develop relatively more quickly and show more variation. In this subsection we discuss some examples of our German data. These examples suggest that the geography of dialect similarity as recorded in the 19th century, in fact, reflects long-term evolutionary processes of region-pair-specific congruencies and past interactions.

---

[8] Linguistic studies on this relationship include Barbujani *et al.* (1996), Dupanloup de Ceuninck *et al.* (2000), and Manni (in press). For a broader discussion, also see the "*linguistic dynamics approach*" developed by Schmidt (2010).

Our first example illustrates the interrelations between dialects and religious similarities. The map on the left in Figure 2 depicts the regional similarities to the dialect spoken in *Waldshut*, a district located in the southwest of Germany (Baden-Württemberg). The reference point *Waldshut* is marked. Warm colors indicate a high, and cold colors a low degree of similarity. The map on the right in Figure 2 zooms in on Baden-Württemberg and compares the spatial pattern of dialect similarity with the religious geography of that area. As is well known, the Reformation of the 16th century resulted in distinct Protestant and Catholic localities in Germany (see also Becker and Woessmann 2009). Protestant areas in the year 1546 are indicated in Figure 2 by a hatching from left to right, whereas the hatching from right to left indicates those areas that were Catholic in 1546 but became Protestant by 1820. Notice that there are only very few such areas, i.e., religious orientation remained remarkably stable over this time span of almost 300 years. This stability is chiefly due to social practice. For example, in earlier times it was uncommon, if not completely unheard of, to marry across religious borders; Protestants marry Protestants, Catholics marry Catholics.[9]

[FIGURE 2 HERE]

The main message conveyed by Figure 2, however, is that the geography of dialect similarity is strikingly similar to religious geography. Waldshut itself was and always remained Catholic, and it can be seen that the dialects of other Catholic districts resemble the one in Waldshut more closely than do the dialects of Protestant districts. This finding aligns itself nicely with an evolutionary perspective on linguistic development. Catholic localities have been in closer contact with other Catholic localities; Protestants have been more in contact with other Protestants. Hence, religious and linguistic similarities have co-evolved, and recent linguistic studies suggest that they do so until today (see Stoeckle 2010).

---

[9] This stability is remarkable in light of the fact that it was not until after the Peace of Westphalia (1648) that a newly-converted ruler became prohibited from forcing his new religion on his subjects, which had been common practice ever since the Peace of Augsburg in 1555 (see Cantoni 2009). Other factors apart from social practice that might have a stabilizing effect on religious orientation include natural boundaries such as the Black Forest or the Rhine, or administrative borders, in this case the border of the archbishopric Freiburg.

Second, language is also reflective of previous migration waves. To illustrate this point, let us consider the example of the *Goslar* district. The map in Figure 3 illustrates the dialect similarity between Goslar (white) and all other German districts.

[FIGURE 3 HERE]

Linguists view the Harz Mountains in Goslar as a *language enclave* in the sense that the dialect spoken there is not similar to dialects spoken in neighboring districts but instead more resembles a dialect spoken about 300 kilometers away in the mountainous *Erzgebirge*, where, in Figure 3, we find an accumulation of warm colors (indicating high similarity). The historical explanation for this phenomenon is the revival of silver mining in the Goslar area between 1520 and 1620, motivating migration to that area by starving miners in Saxony. This 16th-century relationship between the two regions is still visible in dialect data from the late 19th century, which illustrates the degree of inertia inherent to evolutionary processes. An important aspect of pre-modern migration is that it was nearly always a social or *mass* phenomenon (Wiesinger 1983a), and thus much different from current migration, which is strongly based in *individual* motives. With very few exceptions, these mass migrations in Germany ended during the 18th century. Therefore, at the time Wenker conducted his language survey (1879–1888), roughly one and a half centuries had elapsed without such major perturbations.[10] The local cultures and dialects had thus some time to harden.

Third, geographical distance certainly plays a role in dialect similarity. As seen in Figure 2, the districts adjacent to Waldshut tended to have similar dialects. However, we also find districts relatively close to Waldshut that are less similar than districts that are farther away. This suggests that our dialect data contain information that goes beyond what can be explained by mere physical distance, a point made clearly by the Goslar example (Figure 3),

---

[10] The last incident known to us that can be classified, albeit rather broadly, as a *mass migration* occurred between 1749 and 1832. Initially, a rather small community of people from the Palatinate decided to immigrate to America, but ended up as settlers in a region near the city of Kleve. The reason for migrating was hunger caused by a poor harvest. Once settled in that area, other families from the Palatinate followed.

where there is virtually no relationship between geographical distance and dialect similarity. Dialect similarity could, however, still reflect the existence of old trading routes, which, by taking advantage of rivers, natural passages, and forts, historically led to more contact between certain regions. And, indeed, the importance of transport routes for the spatial structuring of language attributes is made evident by the example of the so-called *Rheinstaffel*. Klausmann (1990) notes a difference in linguistic development depending on the topological relation of individual locations to the Rhine river, i.e., dialect similarity may also be influenced by ancient transportation networks.

Last, at the time Wenker collected the data, the German Empire had just been created out of formerly independent territories. These territories had previously been in existence for centuries, and thereby also contributed to linguistic evolution. In fact, dialectologists since the 19th century were aware of the congruencies between the areal distribution of historical territories and language (see Haag 1898; Aubin *et al.* 1926; and, more recently, Barbour and Stevenson 1990). One reason for this persistence may be that the territories tended to encourage internal traffic, and discourage, or at least not improve the means for, travel external to their borders. Hence, communication and exchange between territories was somewhat hindered (Bach 1950:81). From an evolutionary perspective, such limitations can lead to a higher degree of dialect similarities among regions that formerly belonged to the same historic territory.

Summing up, these examples suggest that dialect similarity between regions is higher the more intensive was their interaction and exchange in the course of history. Those influences, such as common religious and historical political borders, distance and the influences of ancient transportation networks, as well as unique historical events and previous migration waves, all left some long-lasting imprints on the local dialects. Dialect similarities between regions are thus correlated with these types of regional congruencies, but are likely to capture also other (and less well measurable) aspects of cultural similarity.

3. A gravity model of current regional migration

In the empirical analysis we investigate to what extent historical dialect differences affect current cross-regional migration flows. In this section, we derive a theoretically grounded gravity equation for migration, which serves as the underlying framework.

3.1. Current regional migration data

We use data on pair-wise gross migration flows for the 439 German districts averaged over the period 2000–2006 as provided by the German Federal Statistical Office.[11]

[TABLE 2 HERE]

Table 2 provides an overview of these data and points out two basic facts about internal migration flows in Germany. First, it indicates that the average annual gross flow between a pair of regions was only seven migrants per 100,000 inhabitants in the district of origin, which implies a total gross emigration rate of only 3% for the typical German district. Gross migration flows in Germany are thus relatively small compared to countries like the US, which suggests that mobility costs are substantial. Second, across all regional pairs, there has been some gross migration in more than 96% of all cases. That is, migration occurs not only from economically poor to rich regions, but also in the other direction. Our gravity model for migration accommodates those features. It takes into account that individuals are heterogeneous and face distance-dependent mobility costs should they decide to move, and it generates two-way gross migration flows which are larger than net flows.

3.2. The model

We consider a simple discrete location choice model with random utility (see Anderson *et al.* 1992). Assume a country that consists of $r = 1, ..., R$ regions, and a mass of heterogeneous individuals indexed by h. Indirect utility of individual h in region r is given by

---

[11] In Germany, every person who changes his or her place of residence is legally required to register at the new residence within at most two weeks (even earlier in some states). The migration data are thus very accurate.

$$V_r^h = u_r + \varepsilon_r^h \tag{1}$$

The term u_r is region-specific and identical across individuals. It includes economic variables like labor and housing market conditions, as well as all regional features that affect everyone in that region in the same way. The term ε_r^h is idiosyncratic for h and r, capturing heterogeneity in the way how region-specific features affect different individuals. For example, this term may include job signals from different regions or occupational shocks pertaining to that particular person h, as well as his or her view on local amenities.

We refer to the term u_r as the *average welfare* of region r, and for our purposes it suffices to think of u_r as being exogenously given. That is, we assume that the average welfare levels do not respond to the individual location decisions.[12] For the idiosyncratic term ε_r^h we assume that the actual matching value between an individual and a region is the realization of a random variable. Specifically, we assume that ε_r^h is distributed i.i.d. across individuals and regions, and we adopt the standard parameterization of a double exponential distribution, $F(x) = \mathbf{Pr}\left(\varepsilon_r^h \leq x\right) = \mathbf{exp[-exp}(-x/\beta - \gamma)\mathbf{]}$, where γ (≈ 0.5572) is the Euler constant, and $\beta > 0$ is a parameter. This distribution has mean zero and variance $\left(\pi^2/6\right) \cdot \beta^2 \approx 1.6449 \cdot \beta^2$. The term β is referred to as the *degree of heterogeneity* of the idiosyncratic shocks.

If cross-regional mobility was costless, the choice probability of some individual h to live in region r would be given as follows (see Murata 2003):

$$P_r = \mathbf{Pr}\left[V_r^h > \max_{j \neq r}\left\{V_j^h\right\}\right] = \frac{\mathbf{exp}\left[u_r/\beta\right]}{\sum_{j=1}^{R}\mathbf{exp}\left[u_j/\beta\right]} \tag{2}$$

With $\beta \to 0$, people make deterministic location choices based only on the average welfare levels. On the other hand, with $\beta \to \infty$, idiosyncratic shocks become extremely heterogeneous

---

[12] See Murata (2003) and Tabuchi and Thisse (2002) for economic geography frameworks that incorporate a random utility model. Those papers consider welfare to be endogenous and solve for long-run equilibria.

and people choose among the *R* regions with equal probability *(1/R)*. We now embed this model into a two-period framework and introduce region-pair-specific mobility costs.

Suppose there is some innate distribution of individuals across regions. The random variables ε_r^h are drawn at the end of the first period. Given the realizations of the ε_r^h and the values of u_r, individuals then choose the location they most prefer during the second period. An individual *h* migrates from the initial location *r* to some other region *s* if the overall utility from living in *s*, net of the mobility costs c_{rs}, exceeds the (net of mobility costs) utility level of all other locations *j*, including the location of birth *r*. Formally, a move from *r* to *s* takes place if $V_s^h - c_{rs} > \max_{j \neq s}\{V_j^h - c_{rj}\}$, with $c_{rr} = 0$ and $c_{rj} \geq 0$ if $j \neq r$. Using (2), the probability of migrating from *r* to *s* is then $P_{rs} = \exp\left[(u_s - c_{rs})/\beta\right]\Big/\sum_{j=1}^{R}\exp\left[(u_j - c_{rj})/\beta\right]$. Aggregating across individuals, the gross migration flow from *r* to *s* is equal to $M_{rs} = P_{rs} \cdot L_r$, where L_r is the population size of the source region. Rearranging $P_{rs} = M_{rs}/L_r$ and taking logs we obtain the following gravity equation:

$$\log\left(M_{rs}/L_r\right) = (u_s - c_{rs})/\beta - \log\left[\sum_{j=1}^{R}\exp[(u_j - c_{rj})/\beta]\right].$$ Notice that the location choice in our framework may involve the decision of an individual to migrate to an area with a lower level of average welfare than in the current source region, as well as the decision to remain in the region of birth. Furthermore, due to the idiosyncratic shocks, the model features parallel gross migration flows from *r* to *s* and from *s* to *r*.

Turning to the specification of the mobility costs, we not only include standard pecuniary costs (for moving furniture, finding accommodation, etc.), which are denoted by d_{rs}. We also incorporate, in the spirit of Sjaastad (1962), non-pecuniary costs of migration, denoted ℓ_{rs}, which capture the psychic costs of moving to a culturally unfamiliar environment.

We assume the following specification: $c_{rs} = a_1 \cdot \log[d_{rs}] + a_2 \cdot \log[\ell_{rs}]$. We can then re-write the gravity equation and arrive at our final equation to be estimated:

$$\log(M_{rs}/L_r) = D_r + D_s + \alpha_1 \cdot \log[d_{rs}] + \alpha_2 \cdot \log[\ell_{rs}] + e_{rs}, \qquad (3)$$

where e_{rs} is a standard error term. Notice that the term $D_r = -\log\left[\sum_{j=1}^{R} \exp((u_j - c_{rj})/\beta)\right]$

varies only at the level of the source region, whereas the term $D_s = u_s/\beta$ varies only at the level of the destination region. These terms will therefore be captured by source and destination area fixed effects in the empirical analysis. The mobility costs vary at the region-pair level. In the empirical analysis, we measure pecuniary mobility costs d_{rs} by the physical distances and travel time across German regions (see Table 1 above for some summary statistics thereof), and the cultural mobility costs ℓ_{rs} by the historical dialect data.

A specification like Equation (3) is standard practice in the gravity literature in international trade. The fixed effects capture all region-specific impact variables in our cross-sectional analysis, such as contemporaneous wages, prices and amenities, as well as institutions and further historically determined regional characteristics. The fixed effects also tackle the problem of interdependent flows in a multi-region economy (Anderson and van Wincoop 2003), since the term D_r captures the overall position of the source area within the entire network of regions (its "multilateral resistance term"). As shown by Feenstra (2004), this specification allows for a consistent estimation of region-pair-specific impacts. The coefficients of interest are the geographical distance elasticity α_1 and, in particular, the elasticity α_2, which measures the impact of dialect similarity on gross migration flows.[13]

---

[13] Since we have $\alpha_2 = a_2/\beta$, we can identify this elasticity up to the unobservable constant $1/\beta$.

4. The effect of dialect similarity on regional migration

We estimate the gravity equation (Equation (3)) by ordinary least squares with origin and destination fixed effects. Table 3 presents the baseline estimation results.[14]

[TABLE 3 HERE]

The results show that dialect similarity has a positive and highly statistically significant effect on gross regional migration flows. When including only dialect similarity without controlling for geographical distance, as in specification 1, we find a sizable cultural mobility costs that impede internal migration in Germany. The (scaled) elasticity of gross migration with respect to dialect similarity is around 2.2.

As illustrated by the examples in Section 2, however, dialect similarity is correlated with geographical distance, which per se has a negative impact on migration flows. To address this issue we first separately study the impact of geographical distance without considering dialect similarity. In specification 2 we use the linear physical distance between the centers of the source and the destination district as our proxy for pecuniary mobility costs. The results show that increasing the physical distance between two regions, all else equal, drives down gross migration flows (the elasticity is about 1.5). In specification 3 we use an alternative distance measure, namely, the travel time by car between any pair of regions (in minutes). This proxy may better capture the true regional accessibilities, as it takes into account irregularities in distances due to natural barriers like rivers or mountains. The results indicate that the elasticity with respect to travel time (1.77) is a bit larger than for physical distance. When including both measures at the same time (as in specification 4), it turns out that most of the negative impact is captured by physical distances, with travel time having some additional small impact. Altogether, these findings are consistent with the previous migration

---

[14] All estimation results reported in this paper refer to the total stock of migrants and regional populations irrespective of age. Results are similar, however, when considering only working-age individuals.

literature that has established a detrimental effect of geographical distance on migration flows (see, e.g., Schwartz 1973; Greenwood 1975).

The important question is whether the positive effect of dialect similarity on migration flows prevails once we control for geographical distance. In specification 5 we simultaneously include dialect similarity and both proxies of pecuniary mobility costs. As can be seen, the coefficient α_2 drops substantially compared to column 1, which is due to the correlation of linguistic and geographical distance. However, even conditional on geographical distance (and origin and destination fixed effects), we find a positive and highly significant effect of dialect similarity on gross migration flows.[15] The estimated elasticity is 0.186. We take this elasticity (column 5 of Table 3) as our benchmark finding throughout this paper.

To set this result into perspective, recall from Table 1 that the standard deviation of our dialect similarity measure is 10.53, which is about 32 percent of the sample mean. Thus, an increase of dialect similarity by one standard deviation (measured at the sample mean) increases gross migration flows between a pair of regions by about 6 percent. We can also separate the direct effect of distance on migration, and the indirect effect operating through dialects. A simple OLS regression across all regional pairs reveals that increasing physical distance by 100km reduces dialect similarity by 4.23 linguistic correspondences on average, i.e., by around 40% of a standard deviation. We can then calculate that decreasing geographical distance by 100km raises gross migration flows by roughly 41% directly, and by another 2-3% because dialects have been more similar.

Columns 6 and 7 of Table 3 address the robustness of our main finding with respect to the estimation method. First, the interpretation of the parameters of log-linear gravity models estimated by linear least squares methods can be misleading in the presence of heteroskedas-

---

[15] In the literature on how genetic similarities affect international trade flows, Giuliano *et al.* (2006) argue that there may actually be no such effects once transport costs across countries are properly controlled for. Our estimation in column 5 takes such issues into account because actual travel time across regions can be thought of as an analogue of actual transport costs for goods.

ticity. To overcome this problem, we estimate the gravity equation by means of a Poisson pseudo-maximum-likelihood (PPML) estimator with Eicker-White robust standard errors, as proposed by Santos-Silva and Tenreyro (2006). Second, previous work in the international trade literature suggests that zero flows can pose problems in the estimation of gravity equations (see Disdier and Head 2008; Helpman *et al.* 2008). As shown above, zero gross migration flows across German districts account for less than 4% of all cases and therefore would appear to be a minor issue. Nevertheless, we tackle this potential problem by employing a two-stage Heckman procedure that uses a non-linear probit equation for selection into migration in the first stage, and then estimates Equation (3) in the second stage.[16] In the PPML estimation (see column 6), the elasticity with respect to dialect similarity is around 0.11 and thus somewhat lower than in the benchmark specification. The Heckman selection model (column 7) yields estimates that are similar to the benchmark.

Table 3 additionally shows that our results are also robust with respect to the linguistic similarity index. We replace the simple count index with the similarity index by Jaccard (1901) in column 8, and with the similarity index by Kulczynksi (1927) in column 9, while returning to ordinary least squares estimation. Regardless of which similarity index we use, our results are similar to the benchmark specification.[17]

In Table 4 we investigate the effect of dialect similarity on migration flows for different types of region pairs. In particular, we divide the 439 German districts into 178 urban and 261 peripheral regions (using the common classification system of the *Federal Office for Building and Regional Planning*), and then distinguish four categories of flows: urban-to-urban (U-U), peripheral-to-peripheral (P-P), urban-to-peripheral (U-P), and peripheral-to-urban (P-U). These four different samples thus permit us to investigate whether the impact of

---

[16] We thus rely on the normality assumption for identification of our second-stage estimates.
[17] See Section 2.1. for more details on these indices. Including any of these similarity indices (or the geographical distance measures) in levels instead of logs does not change our qualitative results. We thus consistently use a logarithmic specification, which allows interpreting those coefficients as elasticities.

dialect similarity on migration decisions is dependent on whether the source and the destination area are rather heterogeneous (as in the U-P and the P-U case) or homogeneous (as in the U-U and the P-P case). The results in Table 4 show, however, that the estimated elasticity is rather similar in all cases.[18]

[TABLE 4 HERE]

To further study effect heterogeneity with respect to different types of migration flows, we have again used the full sample of region pairs and dropped all flows where the destination region is a major city with more than 500,000 inhabitants. Those migrations may be mainly motivated by the career concerns of young workers, who temporarily move because human capital investments are easier in big cities (Glaeser 1999; Peri 2002), whereas cultural affinity may play a lesser role in those cases. Furthermore, to tackle a possible non-linearity in the impact of distance on migration, we have dropped all flows below some minimum distance (50, 100, or 150 km), as individuals may not have to leave their cultural environment over those short distances. We omit the detailed results, but the effect of dialect similarity on migration hardly changes in those additional robustness checks. Finally, we have augmented the benchmark specification with a dummy variable that equals unity if the source and the destination region are not located within the same federal same state (*Bundesland*). Crossing the border of a federal state may systematically increase pecuniary mobility costs, e.g., because of different regulations and laws applicable to various occupational groups (e.g., for teachers and lawyers). It turns out that state borders in fact significantly reduce gross migration flows. The effect of dialect similarity on migration hardly changes, however.

The results reported so far imply that an individual who decides to migrate today, all else equal, prefers destinations with a dialect similar to the one prevalent in his or her source re-

---

[18] The elasticity is lower for the P-P sample, but it should be noted that zero migration flows are concentrated within this group. Although the P-P sample accounts for only 35% of all region pairs, it includes 56.6% of all zero flows. When using the Heckman-estimator for the P-P sample, which specifically deals with the zero flow problem, we obtain a coefficient 0.096[***] (std.error 0.015) which is again closer to the benchmark.

gion more than 120 years ago. Our interpretation of this finding is that dialect differences capture intangible cultural barriers at the regional level, which significantly impede migration in Germany. This interpretation rests on three important conditions:

1. That the historical dialect differences from the 19th century are indeed a relevant measure for contemporaneous dialect (cultural) differences across German regions.

2. That the impact of dialect similarity is not confounded with other factors that drive contemporary migration and that are correlated with the recorded historical dialect patterns. If that were the case, dialect similarity would just be a proxy for "something else" and our estimations would suffer from an omitted variable bias.

3. That the causality is indeed such that dialect (cultural) differences affect migration. The alternative hypothesis would be that migration flows are persistent over time due to some omitted factors, and that they have inter alia shaped the geography of dialects. If that were the case, our estimations would capture a spurious correlation.

In the next section we subsequently address these concerns, thereby shedding light on the economic interpretation of our benchmark result.

5. Robustness checks and causality issues

5.1. Persistence of dialect similarity over time

Are historical linguistic patterns representative for contemporaneous dialect differences across regions? Certainly, dialect use is less common today than it was when Georg Wenker conducted his language survey. That is, *absolute* differences in linguistic attributes are likely to be smaller today than in the 19th century. The ubiquity of modern mass media may be one factor that has facilitated this diffusion; the migration that has occurred during the 20th century, which became an increasingly individual phenomenon, is another factor that possibly

contributed to the local languages. However, as long as these developments do not systematically affect some regions more than others, our empirical approach is still consistent, since the *relative* linguistic differences across regions may be persistent over time and influence migration decisions, even if the absolute language differences have become smaller.

Analyzing the current dialect situation in Germany is difficult, since there is no recent comprehensive language inquiry comparable to the Wenker survey, at least not for the country as a whole. What is available besides rather small case studies, however, is a comprehensive dialect survey for 69 Bavarian districts referring to the time period from the 1980s till today. In this section we explore these data in order to address the time persistence of dialect structures at least for this subset of German regions, and to analyze the effect of contemporaneous dialect similarity on migration within Bavaria using historical dialect similarity as an instrument. The data come from the Bavarian linguistic atlas project (*Bayerischer Sprachatlas*,[19] "BSA"). They allow us to recover today's spatial distribution of the same language characteristics as in the original Wenker survey, and to build a contemporaneous dialect similarity matrix across 69 Bavarian regions in an analogous way as described above.

Figure 4 gives an example. The map on the left illustrates dialect similarity to the district of Passau using the contemporaneous BSA data. The map on the right uses the data from the 19th century and depicts dialect similarity with Passau for the same group of regions.

[FIGURE 4 HERE]

The two maps look virtually identical, which suggests that the geography of dialects remained remarkably stable. This conclusion also holds more generally. The correlation between the historical and the contemporaneous dialect similarity matrix is 0.85, which sug-

---

[19] The BSA data are available from http://www.baydat.uni-wuerzburg.de. At the time of our data collection the data from the northern part of Bavaria were not available yet.

gests that relative linguistic differences are highly persistent over time.[20] Furthermore, comparing the historical and the contemporaneous dialect vectors $\mathbf{i}^r$ for every region, respectively, we find that 81.68% of all entries are identical and less than 18% of the evaluated language characteristics differ between the historical data and the recent data. That is, even in absolute terms there has been no substantial change in the language patterns in Bavaria during this time period of 120 years.[21]

[TABLE 5]

Turning to the econometric analysis, column 1 of Table 5 first reports the results of a fixed effects regression of contemporaneous dialect similarity on regional gross migration flows across the 69 Bavarian regions. Column 2 of Table 5 reports the results of an instrumental variable (IV) specification where we use historical dialect similarity as an instrument for contemporaneous dialect similarity. The F-value of the excluded instrument (6,291.7) indicates that this instrument is indeed highly relevant. Overall, the IV specification supports our main conclusion that dialect similarity positively affects regional migration flows. Comparing the OLS and the IV specification, we find that the estimated elasticity nearly doubles.[22] This suggests that the simple OLS estimator of current dialect similarity on migration flows in column 1 is downward biased.

Such a downward bias might be the result of the linguistic diffusion that occurred during the 120 years since the Wenker survey, i.e., it may capture the impact of such factors where mi-

---

[20] This correlation even increases to 0.88 if one corrects the BSA-matrix for changes in linguistic evaluation techniques since the Wenker survey, and for the few missing values where single language characteristics are not reported for single Bavarian districts.

[21] There are a few additional linguistic case studies on dialect change in Germany, which are mostly dedicated to specific language phenomena. Those studies also note a close correspondence of the current geography of dialects with historical linguistic patterns (e.g., Bellmann 1985:213). In the economic literature, the transmission of cultural values across generations and the stability of cultural patterns over time is further studied in Tabellini (2008) and Benabou and Tirole (2006).

[22] Using the first stage coefficient of historical on contemporaneous dialect similarity (see column 2 of Table 5), we can convert our benchmark elasticity (column 5 in Table 3) into an elasticity of current migration with respect to contemporaneous dialect similarity. This would be 0.276 (= 0.186/0.674), but this calculation has to be taken cautiously as it assumes that the first stage coefficient derived from the Bavarian sample can be generalized to Germany as a whole.

24

gration has influenced dialect similarity. Yet, this diffusion does not affect our IV regression for Bavaria as well as our reduced form estimates of historical dialect similarity on contemporaneous migration flows for Germany as whole. The reduced form estimate yields a consistent proportional estimate for the effect of contemporaneous dialect similarity on migrations flows even in the presence of this linguistic diffusion.[23]

5.2. Omitted region-pair-specific and region-specific characteristics

If the geography of dialects is persistent over time, the question still remains if cultural differences are properly captured by linguistic differences. In that regard, the question remains if there are other persistent (cultural or non-cultural) factors that drive contemporary migration and that are also correlated with historical dialect patterns. In this subsection we deal with this issue of omitted variable bias. Notice that our benchmark estimations are consistent as long as omitted variables are purely *region*-specific, due to the fixed effects specification of our empirical model. A problem would arise, however, if we omit relevant *region-pair*-specific variables that simultaneously affect dialects and migration patterns.

In general, it should be noted that culture is, of course, not restricted to language, but occurs in many other domains such as art, traditions, habits, etc. Yet, a vast literature in anthropology and sociology insists that cultural and linguistic evolution proceed in parallel, and that language is the strongest marker of cultural identity (see, e.g., Cavalli-Sforza 2000; Brewer

---

[23] We have confirmed this conclusion in a further robustness check that contemplates the only major mass migration in Germany during the 20th century, namely the massive inflow of expellees (*Heimatvertriebene*) after World War II. In total, about 12 million expellees were forced to settle within the new borders of Germany and Austria, and they were distributed across regions by the occupation authorities mainly according to the availability of nutrition and housing space. More than 16 percent of all expellees were settled in Bavaria, which accounted for roughly 20 percent of the total Bavarian population in 1950. As the expellees brought along their local dialects (mostly from Silesia and the Sudeten area), this necessarily led to a mix with the prevailing local languages. There is, hence, the concern that linguistic diffusion has disproportionately affected such Bavarian regions were a relatively large share of expellees were forced to settle. We have used data on the regional shares of expellees in 1952, as published by the Federal German Government (*Bundesminister für Vertriebene*), and considered interaction terms of historic dialect similarity with the expellee shares of the destination and origin region as additional instruments in our IV specification. Our second stage coefficient of contemporaneous dialect similarity on migration flows hardly changes when additionally using these instruments.

1991). That is, if two regions share many traditions and habits, this indicates that those regions had numerous interactions in the course of history. This, however, would have left lasting imprints on language structures as well, i.e., similarities within those other cultural domains should be reflected in dialect differences. This implies that linguistic similarity is a relevant and comprehensive measure of deep cultural ties at the regional level in Germany, a view that is consistent with Charles Darwin's claim that languages are the "*best indicator*" for the "*pedigree of mankind*".

It is still important to analyze if really *dialects* impact on migration flows, or rather something else that is correlated with dialects. Specifically, we argued in Section 2 that dialect similarity reveals a spatial pattern that often corresponds to other types of historically determined congruencies between the regions, including religious orientation as indicated by the Waldshut example (Figure 2). We therefore control for differences in religious denominations in 1890, roughly the same time at which the linguistic data were collected. We define a dummy variable that equals unity if the majority of the population in the source region had a different religion than those in the destination region in the late 19th century. If cultural differences between current German regions are manifested mainly along those religious lines, and if dialects simply pick up these persistent effects of religion, we would expect the elasticity of migration with respect to dialect similarity to turn insignificant (or at least to drop substantially) once we include this additional control variable.

Second, the Goslar example (Figure 3) has emphasized the role of industry structures in driving migration flows. If two regions are persistently specialized in the same sectors, there may be constantly more migration between them as human capital is at least partly industry-specific. This industry similarity may continue to drive migration to the extent that local sectoral structures do not change over time. To take this into account, we have constructed an index of historical differences in industrial structures across regional pairs. Data for 24 in-

dustries are available for the year 1907, and are taken from the German workplace and population census. Industrial differences are then calculated as the sum of absolute differences in employment shares over the 24 industries across the different pairs of regions.

Third, another possible confounding factor are former administrative borders, since we emphasized above that the geography of dialect similarity is also correlated with the borders between the territories out of which the German Empire was created (as noted, e.g., by Barbour and Stevenson 1990). We therefore include a dummy that equals unity if the current migration flow extends across a historical administrative border.[24]

Finally, even though we control for contemporaneous travel distances across regions, dialect similarity may also pick up persistent effects of ancient trading routes and transportation networks that have formerly linked particular pairs of regions and, thus, facilitated exchange between them. To cope with this issue, we have calculated a measure for historical travel distance based on a comprehensive German railway system outline by Friedrich List in 1833 (see the data appendix for further details).

[TABLE 6 HERE]

In columns 1 to 4 of Table 6 we control for the new variables separately; they are considered jointly in column 5. The results suggest that there is significantly more current migration between regions with historically different religious denominations, and less migration between regions with stronger differences in their historical industrial structures. The presence of a historical administrative border exerts a negative impact on current migration flows. Historical travel distance does not turn out to be significant when it is considered jointly with contemporaneous travel distance.[25]

---

[24] More specifically, we consider the borders of 38 member states and 4 independent cities that were part of the German Confederation at the time of its foundation in 1815. These borders are a good representation of the politically fragmented environment that prevailed until the German Empire was established.

[25] The positive effect of religious differences on migration may seem counterintuitive. However, it may capture an enduring prosperity difference between Catholic and Protestant areas, which was recognized early on by

The main insight of Table 6, however, is that the elasticity of migration with respect to dialect similarity hardly changes when including those new control variables. This suggests we our results do not suffer from omitted variable bias, since dialects do not merely reflect religious, sectoral or political congruencies across regions. These results are thus consistent with the view that language structures are a comprehensive indicator for cultural similarity.

5.3 Persistence of migration flows: Exploiting the German division and reunification

Turning now to the third estimation issue discussed in Section 4, there is still the concern that migration, rather than culture, is persistent over time and has, inter alia, shaped the geography of dialects. Dialect similarity would then not actually *cause* contemporary migration, and our estimations would capture a spurious correlation.

Such persistence in migration can be due to network effects, which result in a clustering of migrants from particular source regions at particular destinations. To illustrate this, suppose that at the time Georg Wenker collected the linguistic data, there was already a previously established migration connection between two regions *r* and *s*. Say, families have constantly migrated across those two regions, because the regions traditionally offered suitable jobs due to similar industrial structures, or because of existing networks of social contacts.[26] To the extent that these network effects remain in operation, they constantly influence migration decisions, which in turn would lead to a correlation of region-pair-specific migration flows over time. If this is so, the prediction could be that dialect distance slowly disappears between those regions *r* and *s* experiencing high migration exchange.

Max Weber and studied further by Becker and Woessmann (2009). Moreover, we find that the historical border dummy turns insignificant when we add current administrative borders in the same way as described in Section 4. This suggests that current and historical borders overlap, so that the historical borders partly capture the negative impact of federal state borders on migration that operates via an increase in pecuniary mobility costs. Finally, instead of historical differences we could also control for contemporaneous differences in industrial structures. Our main result for the impact of dialect similarity would not change.

[26] Network effects in migration are extensively studied both theoretically (Carrington *et al.* 1996) and empirically (e.g., Munshi 2003).

To address this issue, we exploit a quasi-natural experiment in German history. Since the construction of the Berlin Wall in 1961, migration flows between East and West Germany were exogenously cut off.[27] In other words, persistent migration flows that might have caused ongoing dialect assimilation between particular East and West German regions, as well as other types of cross-border contacts, were almost completely interrupted for a considerable time span between the Wenker survey and our current migration data.

When migration between the East and the West became again possible after 1989, the pre-existing migration networks have thus not been in operation for almost 30 years. To the extent that these networks have no "memory", we would thus not expect to see a continuation of the persistence in migration flows across particular pairs of regions that existed prior to the division of Germany.[28] On the other hand, dialect similarity does have such a memory, as emphasized by our various examples, and is likely to have survived the division.

Put differently, if our baseline findings only reflect the persistence of migration flows, we would expect to find no (or at least substantially lower) effects of dialect similarity on contemporary migration flows within a sample of migration flows across the inner German border only. By contrast, if we still find a positive effect of dialect similarity on contemporary migration flows for these cases, such would suggest that cultural identity at the regional level actually does affect migration decisions.

[TABLE 7 HERE]

Table 7 shows the results for the East-West and the West-East samples of region pairs and, indeed, the coefficient of language similarity is still significantly positive and of similar

---

[27] The division and reunification of Germany is used as a quasi-natural experiment by Redding and Sturm (2008), who show that the decline of West German cities near the inner German border can be attributed to the loss of market access to the neighboring East German areas after the division of Germany.

[28] Networks based on similar industry structures were likely to be seriously damaged, because the central planning in the East German economy has led to major perturbations in the spatial distribution of industries (see, e.g., Heinzmann and Karrasch, 1990). The survival of social networks was put under strain by the severe restrictions on personal interactions and face-to-face communications across the inner German border. According to literature on social networks, such personal interactions are essential for the prevalence of a network, which in the absence of communication exhibits a rapid depreciation (see, e.g., Glaeser *et al.* 2002).

magnitude as in the benchmark specification. These results are thus much more in line with an effect of persistent cultural similarity on migration flows, rather than with the opposite causality of persistence in migration flows.

5.4 Persistence of migration flows: First-nature geographical features

Finally, in the last step of the analysis, we investigate another possible source of persistence in migration flows that may have caused the geography of dialects. Specifically, there may be deep regional differences that have persistently driven migration flows over the course of history and, thereby, also linguistic development. In particular, think of first-nature geographical features which have determined the economic prosperity of the regions over the very long run. Salient candidates are indicators of a region's suitability for agriculture and forestry, all of which were major sources of wealth before the Industrial Revolution.

As argued by Combes *et al.* (2010), *soil* characteristics can be regarded as a major determinant of local labor demand in an agrarian society. Accordingly, geological indicators for the suitability of the soil for agriculture and forestry should provide a meaningful insight into the distribution of regional wealth before the heyday of industrialization. These soil characteristics should then be related to ancient migration patterns. As regions with good soil tended to be economically prosperous, they were likely to attract mass migration waves, particularly from areas with bad soil characteristics. A similar point can be made for the *slope* of a region, which is also likely to have influenced agricultural productivity, hence regional prosperity, in former times. Slope may have had another effect on ancient migration patterns – transport routes probably avoided large differences in steepness or ruggedness.

If these very basic geological factors have affected migration waves over the very long run, they could also have influenced the spatial pattern of dialects in Germany. Specifically, the smaller the difference in soil quality and the larger the slope difference between two regions,

the lower the probability that local populations interacted very often. This, in turn, may have resulted in less similar dialects between such regions. To the extent that these geological features still affect current regional migration, our estimations may be capturing a spurious correlation between dialect similarity and migration flows.

The fixed effects specification of the gravity model should, in principle, take into account this potential problem. Consider a region with very favorable geographical features. The resulting pull effects on migration into that region, which have persistently occurred across time and may still occur today, should be captured in the estimation: The fixed effects should level all actual differences in economic prosperity between the origin and the destination, regardless of whether these differences have their origin in history or are the result of current developments. However, to complement this approach, we again create different subsamples of regions that limit the degree of heterogeneity of the respective source and destination areas. For pairs of regions with similar soil and slope characteristics, we may expect very long-run push and pull effects to matter relatively little. This may have led to few cross-regional contacts and therefore to little dialect assimilation over the very long run. In other words, if we find that dialect similarity matters for current migration also for these homogeneous pairs of regions, then a long-run persistence of migration flows is unlikely to be reason. Such a finding would rather suggest that there is actually a genuine effect of cultural similarity on migration decisions.

To address this issue, we sort regions into those with "good" soil and those with "bad" soil. Good soil is suitable and imposes no limitations for agriculture, whereas bad soil imposes such limits because the soil is overly gravelly, stony, or lithic.[29] Using this classification scheme, we can create subsamples of regional pairs and separately study migration flows for

---

[29] We are indebted to Gilles Duranton for providing the data for these indicators (see the Appendix and Combes *et al.* for a more detailed description). To use current indicators of soil quality we need to assume that soil characteristics have not changed during the past centuries, and there are good reasons to believe that this condition is met by our binary distinction between good and bad soil. We also tried a variety of other indicators related to the climate and soil of a region, but this did not crucially affect our empirical results.

cases where both the source and the destination area have good soil, where the source has bad but the destination has good soil, etc. A similar approach is adopted to distinguish between regions with different slope characteristics. Slope is measured as the difference between the maximum and minimum elevation in meters within a region. We can then classify "steep" (above average) and "flat" (below average) regions and create appropriate samples of regional pairs. The results of our gravity estimation for these samples of regional pairs are reported in Table 8a and 8b, respectively.

[TABLES 8a AND 8b HERE]

As can be seen, the results are qualitatively similar for all the considered samples. That is, even for those cases where source and destination area are relatively homogeneous in their geographical features, we find a positive and significant impact of dialect (cultural) similarity on current gross migration flows (see columns 1 and 2 of Tables 8a and 8b).

Results do not change when we additionally control for the similarity of the industry structure in the source and destination area and the historical connectedness of regions via ancient trading routes (see Section 5.2), which is a sensible exercise since first-nature geographical features may also have influenced industry compositions and transport networks and thereby have an indirect effect on current migration flows. Results also do not change when considering an encompassing specification where all control variables and all dummy variables used for sample splits that we have considered so far are included all at once.[30]

5. Conclusion

In this paper, we have evaluated detailed linguistic micro-data from the 19th century on the intra-national variation of phonological and grammatical attributes within the German lan-

---

[30] Specifically, when using the same specification as in column 5 of Table 6, and including additional dummy variables for region pairs along the dimensions urban/periphery, East/West Germany, good/bad soil and flat/steep slope, we obtain an elasticity of migration with respect to historical dialect similarity equal to 0.084*** (std.error 0.026). This is lower than the benchmark estimate, but still highly statistically significant.

guage. In our empirical analysis we find an economically meaningful effect of historical dialect similarity on current regional migration flows. This finding implies that there are intangible cultural borders within a country that impede economic exchange across its regions. These intangible regional borders in Germany are enormously persistent over time; they have been developed over centuries, and so they are likely to be there also tomorrow.

Even on a low geographical level people seem to be unwilling to move to culturally unfamiliar environments. The average Bavarian will not easily move to Saxony, nor vice versa, unless he or she is compensated by considerably better economic prospects or job opportunities in the other region. The existence of cultural borders thus clearly limits mobility across local labor markets and, thus, the integration of the national labor market.

It is beyond the scope of this paper to discuss whether it is possible, or desirable, to downsize such borders. Policy initiatives in the European Union aiming for a preservation of regional languages tend to suggest that there is currently no interest in cultural equalization, but rather that linguistic diversity is perceived as valuable for a society. It is thus a natural extension for future research to explore the welfare consequences of cultural differences at a low geographical level in greater detail.

References

Alesina, A., and E. La Ferrara (2005). Ethnic Diversity and Economic Performance. *Journal of Economic Literature*, 43(3), 762–800.

Anderson, J. E., and E. van Wincoop (2003). Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review* 93(1), 170–192.

Anderson, S. P., A. de Palma, and J.-F. Thisse (1992). *Discrete Choice Theory of Product Differentiation*. Cambridge, MA: MIT Press.

Aubin, H., T. Frings, and J. Müller (1926). *Kulturströmungen und Kulturprovinzen in den Rheinlanden. Geschichte, Sprache, Volkskunde*. Bonn: Röhrscheid.

Bach, A. (1950). *Deutsche Mundartforschung. Ihre Wege, Ergebnisse und Aufgaben*. 2nd edition. Heidelberg: Winter.

Barbujani, G., M. Stenico, L. Excoffier, and L. Nigro (1996). Mitochondrial DNA Sequence Variation Across Linguistic and Geographic Boundaries in Italy. *Human Biology* 68(2), 201–215.

Barbour, S., and P. Stevenson (1990). *Variation in German. A Critical Approach to German Sociolinguistics*. Cambridge: Cambridge Univ. Press.

Becker, S. O., and L. Woessmann (2009). Was Weber Wrong? A Human Capital Theory of Protestant Economic History. *Quarterly Journal of Economics* 124(2), 531–596.

Bellmann, G. (1985). Substandard als Regionalsprache. In G.Stötzel (ed.) *Germanistik – Forschungsstand und Perspektiven*. Part 1: Germanistische Sprachwissenschaft. Berlin, New York: de Gruyter, 211–218.

Benabou, R., and J. Tirole (2006). Belief in a Just World and Redistributive Politics. *Quarterly Journal of Economics* 121, 699–746.

Brewer, M. B. (1991). The Social Self: On Being the Same and Different at the Same Time. *Personal and Social Psychology Bulletin* 17(5), 475–482.

Cantoni, D. (2009). The Economic Effects of the Protestant Reformation: Testing the Weber Hypothesis in the German Lands. Harvard Univ. Working Paper.

Carrington, W.J., E. Detragiache, and T. Vishwanath (1996). Migration with Endogenous Moving Costs. *American Economic Review* 86(4), 909-930.

Cavalli-Sforza, L. L. (2000). *Genes, Peoples, and Languages*. London: Penguin.

Chambers, J. K., and P.Trudgill (1998). *Dialectology.* 2nd edition. Cambridge: Cambridge Univ. Press.

Chiswick, B. R. (1991). Speaking, Reading, and Earnings among Low-Skilled Immigrants. *Journal of Labor Economics* 9(2), 149-170.

Combes, P.-P., G. Duranton, L. Gobillon, and S. Roux (2010). Estimating Agglomeration Economies with History, Geology, and Worker Effects. In E. Glaeser (ed.) *The Economics of Agglomeration*, Chicago: Univ. of Chicago Press.

Crystal, D. (1987). *The Cambridge Encyclopedia of Language*. Cambridge: Cambridge Univ. Press.

Desmet, K., M. Le Breton, I. Ortuno-Ortin, and S. Weber (2009). The Stability and Breakup of Nations: A Quantitative Analysis. Unpublished Manuscript, Universidad Carlos III Madrid.

Disdier, A.-C., and K. Head (2008). The Puzzling Persistence of the Distance Effect on Bilateral Trade. *Review of Economics and Statistics* 90(1), 37–48.

Dupanloup de Ceuninck, I., S. Schneider, A. Langaney, and L. Excoffier (2000). Inferring the Impact of Linguistic Boundaries on Population Differentiation: Application to the Afro-Asiatic-Indo-European Case. *European Journal of Human Genetics* 8(10), 750–756.

Feenstra, R. (2004). *Advanced International Trade*. Princeton: Princeton Univ. Press.

Ginsburgh, V. and S. Weber (2010). The Economics of Linguistic Diversity – How Many Languages Make Sense?. Forthcoming: Princeton: Princeton Univ. Press.

Giuliano, R., A. Spilimbergo, and G. Tonon (2006). Genetic, Cultural and Geographical Distances. IZA Working Paper 2229.

Glaeser, E. L. (1999). Learning in cities. *Journal of Urban Economics* 46(2), 254-277.

Glaeser, E. L., D. Laibson, and B. Sacerdote (2002). An Economic Approach to Social Capital. *Economic Journal* 112(483), 437–458.

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

Grogger, J. (2010). Speech Patterns and Racial Wage Inequality. Forthcoming: *Journal of Human Resources*

Guiso, L., P. Sapienza, and L. Zingales (2009). Cultural Biases in Economic Exchange? *Quarterly Journal of Economics* 124(3), 1095–1131.

Haag, K. (1898). *Die Mundarten des oberen Neckar- und Donaulandes*. Reutlingen: Hutzler.

Heinzmann, J. and P. Karrasch (1990). Spatial Patterns of Technological Change in the G.D.R. *Geoforum* 21(2), 193-201.

Helpman, E., M. Melitz, and Y. Rubinstein (2008). Estimating Trade Flows: Trading Partners and Trading Volumes. *Quarterly Journal of Economics* 123(2), 441–487.

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

Jaccard, P. (1901). Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bulletin del la Société Vaudoise des Sciences Naturelles* 37, 547–579.

Klausmann, H. (1990). Staatsgrenze als Sprachgrenze? Zur Entstehung einer neuen Wort- und Sprachgebrauchsgrenze am Oberrhein. In: Kremer, L and H. Niebaum (eds.) *Grenzdialekte. Studien zur Entwicklung kontinentalwestgermanischer Dialektkontinua*. Hildesheim, Zürich, New York: Olms, 193–210.

Kulczynski, S. (1927). Classe des Sciences Mathématiques et Naturelles. *Bulletin International de l'Academie Polonaise des Sciences et des Lettres*, Série B, 57-203.

Lameli, A. (2008). Was Wenker noch zu sagen hatte...Die unbekannten Teile des 'Sprachatlas des deutschen Reichs'. *Zeitschrift für Dialektologie und Linguistik* 75(3), 255–281.

Lazear, E. P. (1999). Culture and Language. *Journal of Political Economy* 107(6), S95–S126.

Manni, F. (in press). Sprachraum und Genetics. In: A. Lameli, R. Kehrein, and S. Rabanus (eds.) *Language and Space. Vol. 2: Language Mapping*. Berlin, New York: de Gruyter.

Melitz, J. (2008). Language and Foreign Trade. *European Economic Review* 52(4), 667–699.

Munshi, K. (2003). Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market. *Quarterly Journal of Economics* 118(2), 549-600.

Murata, Y. (2003). Product Diversity, Taste Heterogeneity, and Geographic Distribution of Economic Activities: Market vs. Non-Market Interactions. *Journal of Urban Economics* 53(1), 126–144.

Peri, G., (2002). Young Workers, Learning and Agglomerations. *Journal of Urban Economics* 52(3), 582 – 607.

Rauch, J. (1999). Networks Versus Markets in International Trade. *Journal of International Economics* 48(1), 7–35.

Rauch, J., and V. Trindade (2002). Ethnic Chinese Networks in International Trade. *Review of Economics and Statistics* 84(1), 116–130.

Redding, S. J., and D. M. Sturm (2008). The Costs of Remoteness: Evidence from German Division and Reunification. *American Economic Review* 98(5), 1766-97.

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

Schmidt, J.E. (2010). Linguistic Dynamics Approach. In: Auer, P. and J.E.Schmidt (eds.) *Language and Space. An International Handbook. Vol. 1: Theories and Methods*. Berlin, New York: Mouton de Gruyter, 201-225.

Schwartz, A. (1973). Interpreting the Effect of Distance on Migration. *Journal of Political Economy* 81(5), 1153–1169.

Sjaastad, L. A. (1962). The Costs and Returns of Human Migration. *Journal of Political Economy* 70(5), 80–93.

Spolaore, E., and R. Wacziarg (2009). The Diffusion of Development. *Quarterly Journal of Economics* 124(2), 469–529.

Steger, H., E. Gabriel, and V. Schupp (1989). *Südwestdeutscher Sprachatlas*. Marburg: Elwert.

Stoeckle, P. (2010): Subjektive Dialektgrenzen im alemannischen Dreilandereck, In: Hundt, M., C.A. Anders, and A. Lasch (eds.) *perceptual dialectology – Neue Wege der Dialektologie*. Berlin, New York: de Gruyter, 291-315.

Tabellini, G. (2008). The Scope of Cooperation: Norms and Incentives. Quarterly Journal of Economics 123(3), 905–950.

Tabellini, G. (2010). Culture and Institutions: Economic Development in the Regions of Europe. *Journal of the European Economic Association* 8(4), 677–716.

Tabuchi,T., and J.-F. Thisse (2002). Taste Heterogeneity, Labor Mobility and Economic Geography. *Journal of Development Economics* 69(1), 155–177.

Wiesinger, P. (1983a). Deutsche Dialektgebiete außerhalb des deutschen Sprachgebiets: Mittel-, Südost- und Osteuropa, in: W. Besch et al. (eds.) *Dialektologie. Ein Handbuch zur deutschen und allgemeinen Dialektforschung*. Zweiter Halbbd. Berlin, New York: de Gruyter, 900–930.

Wiesinger, P. (1983b). Die Einteilung der deutschen Dialekte. In: W. Besch et al. (eds.) *Dialektologie. Ein Handbuch zur deutschen und allgemeinen Dialektforschung*. Zweiter Halbbd. Berlin, New York: de Gruyter, 807–900.

Wrede, F., W. Mitzka, and B. Martin (1927–1956). *Deutscher Sprachatlas. Auf Grund des von Georg Wenker begründeten Sprachatlas des Deutschen Reichs*. Marburg: Elwert.

Table 1: Descriptive Statistics of the Similarity and Distance Measures

	Obs.	Mean	Std. Dev.	Min	Max
Dialect Similarity	96,141	32.60	10.53	11	66
Dialect Similarity (Jaccard)	96,141	0.344	0.156	0.091	1
Dialect Similarity (Kulczynski)	96,141	0.494	0.160	0.167	1
Geographical Distance (in kilometers)	96,141	309.69	152.33	1.07	845.32
Travel Distance (in minutes)	96,141	263.33	113.44	8.80	683.31

Notes: Descriptive Statistics are provided for 96,141 pairs of regions r and s (r ≠ s and r=s=439). See section 2.1. for the definition of the linguistic similarity indices.

Table 2: Descriptive Statistics of Gross Migration Flows, Average 2000–2006

	Mean of M_{rs}/L_r (per 100,000 inhabitants)	Mean of all positive M_{rs}/L_r (per 100,000 inhabitants)	Share of district pairs with $M_{rs}/L_r > 0$
German inhabitants, entire population	7.11	7.35	96.75%
German inhabitants, working-age population (18–65)	8.84	9.21	96.04%

Notes: Means are calculated across 192,282 observations for migration flows from every region r to s (r ≠ s and r=s=439). The number of positive observations is 186,025 (184,667) for the entire population (working-age population).

Table 3: Baseline Results—FE-OLS Regressions

	(1) $\ln(M_{rs}/L_r)$ OLS	(2) $\ln(M_{rs}/L_r)$ OLS	(3) $\ln(M_{rs}/L_r)$ OLS	(4) $\ln(M_{rs}/L_r)$ OLS	(5) $\ln(M_{rs}/L_r)$ OLS	(6) M_{rs} Poisson	(7) $\ln(M_{rs}/L_r)$ Heckman	(8) $\ln(M_{rs}/L_r)$ OLS	(9) $\ln(M_{rs}/L_r)$ OLS
Dialect Similarity	2.209*** (0.031)	-	-	-	0.186*** (0.025)	0.118** (0.046)	0.204*** (0.008)	-	-
Dialect Similarity (Jaccard)	-	-	-	-	-	-	-	0.175*** (0.019)	-
Dialect Similarity (Kulczynski)	-	-	-	-	-	-	-	-	0.186*** (0.025)
Geographical Distance	-	-1.493*** (0.012)	-	-1.263*** (0.036)	-1.262*** (0.035)	-1.471*** (0.028)	-1.263*** (0.013)	-1.257*** (0.035)	-1.262*** (0.035)
Travel Distance	-	-	-1.773*** (0.014)	-0.283*** (0.029)	-0.200*** (0.046)	-0.460*** (0.037)	-0.224*** (0.016)	-0.181*** (0.045)	-0.200*** (0.046)
Mills Lambda	-	-	-	-	-	-	0.533*** (0.018)	-	-
R²	0.558	0.744	0.731	0.744	0.745	-	-	0.745	0.745
Pseudo R²	-	-	-	-	-	0.196	-	-	-
Cens. Obs.	-	-	-	-	-	-	6,257	-	-
N	186,025	186,025	186,025	186,025	186,025	192,282	192,282	186,025	186,025

Notes: This table reports estimation results with fixed effects for both origin region r and target region s. In Columns (1)–(7) language similarity is measured by a count index, while Column (8) applies Jaccard's similarity index and Column (9) applies Kulczynski's similarity index. Column (6) reports a Poisson regression of geographical distance and language similarity on the number of German migrants from region r to s, M_{rs}. Column (7) reports the results from a Heckman selection model. In this specification, a first-stage selection considers the probability of a zero flow of migrants between region r and s. Zero flows drop out except in specifications (6) and (7). Geographical distance, travel time, and dialect similarity are in logs in all specifications. Robust standard errors are reported in parentheses.

*** statistically significant at the 1% level; ** statistically significant at the 5% level; * statistically significant at the 10% level.

Table 4: Subsamples: Urban-Periphery

|	(1)	(3)	(3)	(4)
	$\ln\left(M_{rs}/L_r\right)$	$\ln\left(M_{rs}/L_r\right)$	$\ln\left(M_{rs}/L_r\right)$	$\ln\left(M_{rs}/L_r\right)$
	UU	PP	UP	PU
Dialect Similarity	0.180***	0.065*	0.257***	0.208***
	(0.040)	(0.034)	(0.040)	(0.037)
Geographical Distance	-1.632***	-1.211***	-1.037***	-1.049***
	(0.059)	(0.054)	(0.061)	(0.060)
Travel Distance	0.340***	-0.486***	-0.351***	-0.362***
	(0.073)	(0.068)	(0.081)	(0.074)
R²	0.834	0.678	0.710	0.759
N	31,174	64,308	45,176	45,367

Notes: *This table reports OLS results with fixed effects for both origin region r and target region s. In column (1) we consider migration flows where the origin and destination are both "urban" regions. In column (2) we consider migration flows where the origin and destination are both "peripheral" regions. In column (3) we consider urban-to-peripheral, and in column (4) we consider peripheral-to-urban migration flows. "Urban" regions are defined as regional types 1–5 in the classification system of the German Federal Board for Regional Planning (BBR). "Peripheral" areas are defined as regional types 6–9.Robust standard errors are reported in parentheses.*
*** statistically significant at the 1% level; ** statistically significant at the 5% level; * statistically significant at the 10% level.

Table 5: Contemporaneous dialect similarity and migration – Instrumental variable regression

	(1) $\ln\left(M_{rs}/L_r\right)$ OLS	(2) $\ln\left(M_{rs}/L_r\right)$ IV
Contemporaneous Dialect Similarity	0.098*** (0.016)	0.184*** (0.011)
Geographical Distance	-1.654*** (0.087)	-1.525*** (0.052)
Travel Distance	-0.115 (0.105)	-0.071 (0.065)
Historical Dialect Similarity	--	0.674*** (0.009)
F-Test of Excluded Instrument	--	6,291.7
(Second Stage) R²	0.877	0.874
N	4,549	4,549

Notes: This table reports OLS and IV results for the effect of contemporaneous dialect similarity on gross migration flows between 69 Bavarian regions.
Regressions include fixed effects for both origin region r and target region s. In Column (2), historical dialect similarity is used as instrument for contemporaneous dialect similarity.
Robust standard errors are reported in parentheses.
*** statistically significant at the 1% level; ** statistically significant at the 5% level; * statistically significant at the 10% level.

Table 6: Region-Pair-Specific Differences

	(1)	(2)	(3)	(4)	(5)
	$\ln\left(M_{rs}/L_r\right)$	$\ln\left(M_{rs}/L_r\right)$	$\ln\left(M_{rs}/L_r\right)$	$\ln\left(M_{rs}/L_r\right)$	$\ln\left(M_{rs}/L_r\right)$
Dialect Similarity	0.184*** (0.025)	0.170*** (0.025)	0.132*** (0.025)	0.186*** (0.025)	0.114*** (0.025)
Geographical Distance	-1.265*** (0.035)	-1.254*** (0.035)	-1.245*** (0.035)	-1.262*** (0.035)	-1.243*** (0.035)
Travel Distance	-0.201*** (0.046)	-0.193*** (0.046)	-0.161*** (0.045)	-0.200*** (0.046)	-0.157*** (0.045)
Religious Borders	0.018 (0.011)	-	-	-	0.026** (0.010)
Historical Differences in Industrial Structures	-	-1.002*** (0.066)	-	-	-0.934*** (0.065)
Historical Political Borders	-	-	-0.300*** (0.018)	-	-0.295*** (0.018)
Historical Travel Distance	-	-	-	0.001 (0.002)	0.003 (0.002)
R²	0.745	0.746	0.749	0.745	0.751
N	186,025	186,025	186,025	186,025	186,025

Notes: This table reports OLS results with fixed effects for both origin region r and target region s. In column (1) we control for differences in religious denominations in 1890 by including a dummy variable that equals unity if the majority of the population in the source region had a different religion than those in the destination region. In column (2) we include a measure for industry differences in 1907 meassured as sum of differences in the employment shares of 24 industries. In column (3) we include a dummy that equals unity if the current migration flow extends across a historical administrative border between 38 member states and 4 independent cities that were part of the German Confederation at the time of its foundation in 1815.In column (4) we control for historic travel distance measured on a scale from 0 (no distance) to 7 (highest distance). In column (5), we finally include all controls at the same time.
Robust standard errors are reported in parentheses.*** statistically significant at the 1% level; ** statistically significant at the 5% level; * statistically significant at the 10% level.

Table 7: Subsample: East-West

	(1)	(2)	(3)
	$\ln\left(M_{rs}/L_r\right)$	$\ln\left(M_{rs}/L_r\right)$	$\ln\left(M_{rs}/L_r\right)$
	East-West	*West-East*	*East-West and West-East*
Dialect Similarity	0.213***	0.160***	0.187***
	(0.036)	(0.033)	(0.024)
Geographical Distance	-1.580***	-1.443***	-1.513***
	(0.067)	(0.073)	(0.050)
Travel Distance	-0.507***	-0.508***	-0.507***
	(0.082)	(0.073)	(0.056)
R²	0.708	0.534	0.633
N	35,581	34,023	69,604

Notes: *This table reports OLS results with fixed effects for both origin region r and target region s. In column (1) we consider migration flows where the origin is located in former East Germany and the destination is located in former West Germany. In column (2) we consider migration flows where the origin is located in former West Germany and the destination is located in former East Germany. In column (3) we pool East-to-West and West-to-East migration flows. Robust standard errors are reported in parentheses.*
*** statistically significant at the 1% level; ** statistically significant at the 5% level; * statistically significant at the 10% level.

Table 8a: Subsample: Soil Quality

	(1)	(2)	(3)	(4)
	$\ln\left(M_{rs}/L_r\right)$	$\ln\left(M_{rs}/L_r\right)$	$\ln\left(M_{rs}/L_r\right)$	$\ln\left(M_{rs}/L_r\right)$
	Good-Good	Bad-Bad	Good-Bad	Bad-Good
Dialect Similarity	0.179***	0.099*	0.223***	0.194***
	(0.032)	(0.056)	(0.028)	(0.052)
Geographical Distance	-1.431***	-1.127***	-1.123***	-1.195***
	(0.048)	(0.071)	(0.056)	(0.070)
Travel Distance	0.004	-0.510***	-0.333***	-0.259***
	(0.063)	(0.090)	(0.068)	(0.091)
R²	0.748	0.760	0.751	0.727
N	71,836	26,529	43,803	43,857

Notes: This table reports OLS results with fixed effects for both origin region r and target region s. In column (1) we consider migration flows where the origin and destination both have good soil quality. In column (2) we consider migration flows where the origin and destination both have bad soil quality. In column (3) we consider migration flows from regions with good to regions with bad soil quality, and in column (4) we consider migration flows from regions with bad to regions with good soil quality. "Good soil quality" refers to regions with no limitations to agricultural use according to the European Soil Database (esdb) compiled by the European Soil Data Centre. "Bad soil quality" refers to regions with one ore more limitations to agricultural use. Robust standard errors are reported in parentheses.
*** statistically significant at the 1% level; ** statistically significant at the 5% level; * statistically significant at the 10% level.

Table 8b: Subsample: Slope

	(1)	(2)	(3)	(4)
	$\ln\left(M_{rs}/L_r\right)$	$\ln\left(M_{rs}/L_r\right)$	$\ln\left(M_{rs}/L_r\right)$	$\ln\left(M_{rs}/L_r\right)$
	Steep-Steep	*Flat-Flat*	*Steep-Flat*	*Flat-Steep*
Dialect Similarity	0.056	0.246***	0.298***	0.304***
	(0.036)	(0.050)	(0.041)	(0.044)
Geographical Distance	-1.359***	-1.335***	-1.110***	-1.094***
	(0.042)	(0.073)	(0.083)	(0.072)
Travel Distance	-0.281***	-0.286***	-0.284***	-0.266***
	(0.057)	(0.096)	(0.101)	(0.087)
R²	0.734	0.832	0.750	0.717
N	88,628	18,236	39,250	39,911

Notes: This table reports OLS results with fixed effects for both origin region r and target region s. In column (1) we consider migration flows where the origin and destination both are steep regions. In column (2) we consider migration flows where the origin and destination both are flat regions. In column (3) we consider migration flows from regions with steep slope to regions with good slope, and in column (4) we consider migration flows from regions with flat slope to regions with good slope. For each region, slope is measured as the difference between the maximum and minimum elevation in meters. We can then classify a region ith above-average slope as "steep", and with below-average slope as "flat". Robust standard errors are reported in parentheses.
*** statistically significant at the 1% level; ** statistically significant at the 5% level; * statistically significant at the 10% level.

Figure 1a: Exemplary Questionnaire of the Language Survey

**Figure 1b: Exemplary Hand-Drawn Map by Georg Wenker**
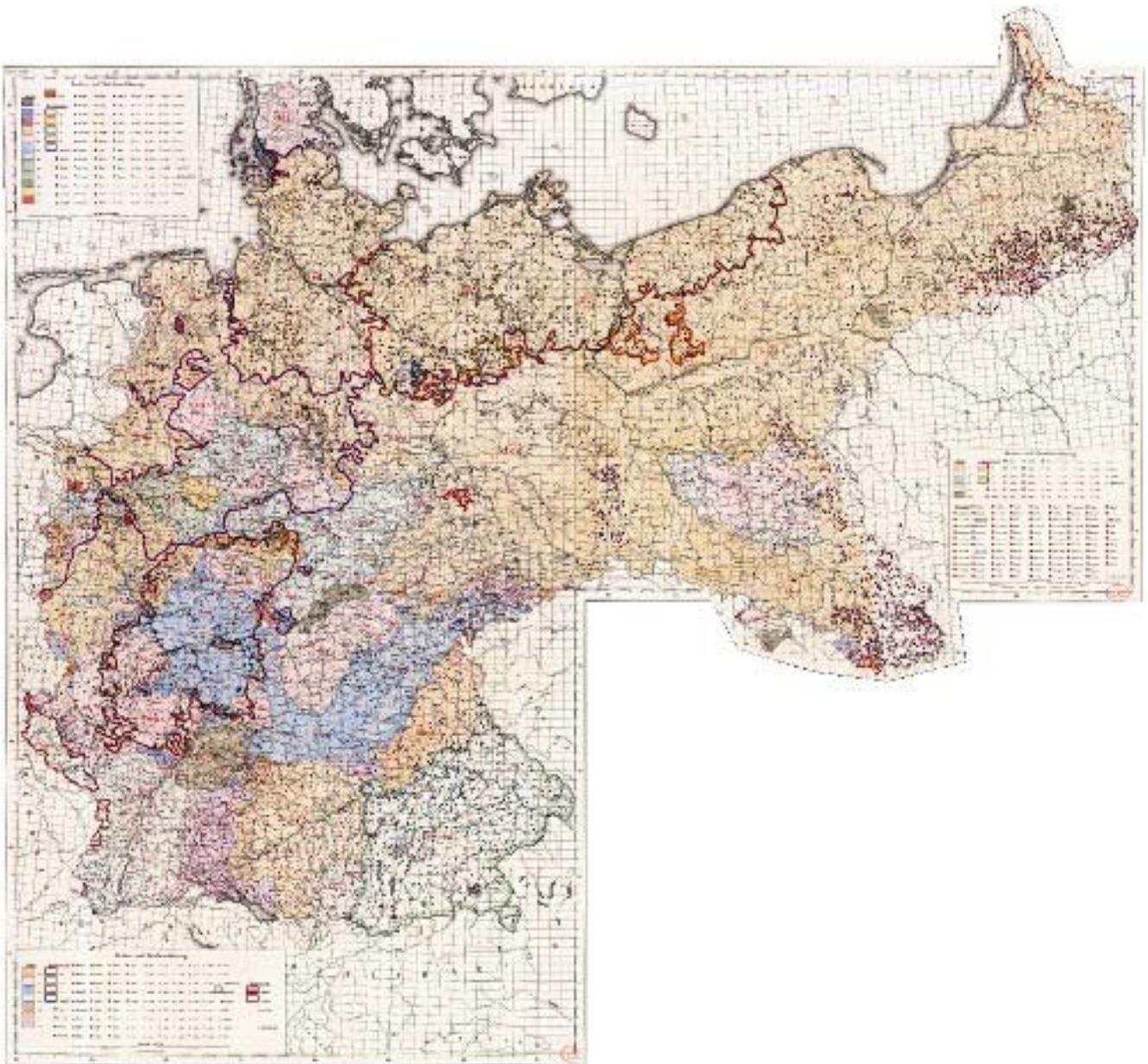
Figure 2: Distribution of Religious Denomination in Southern Germany
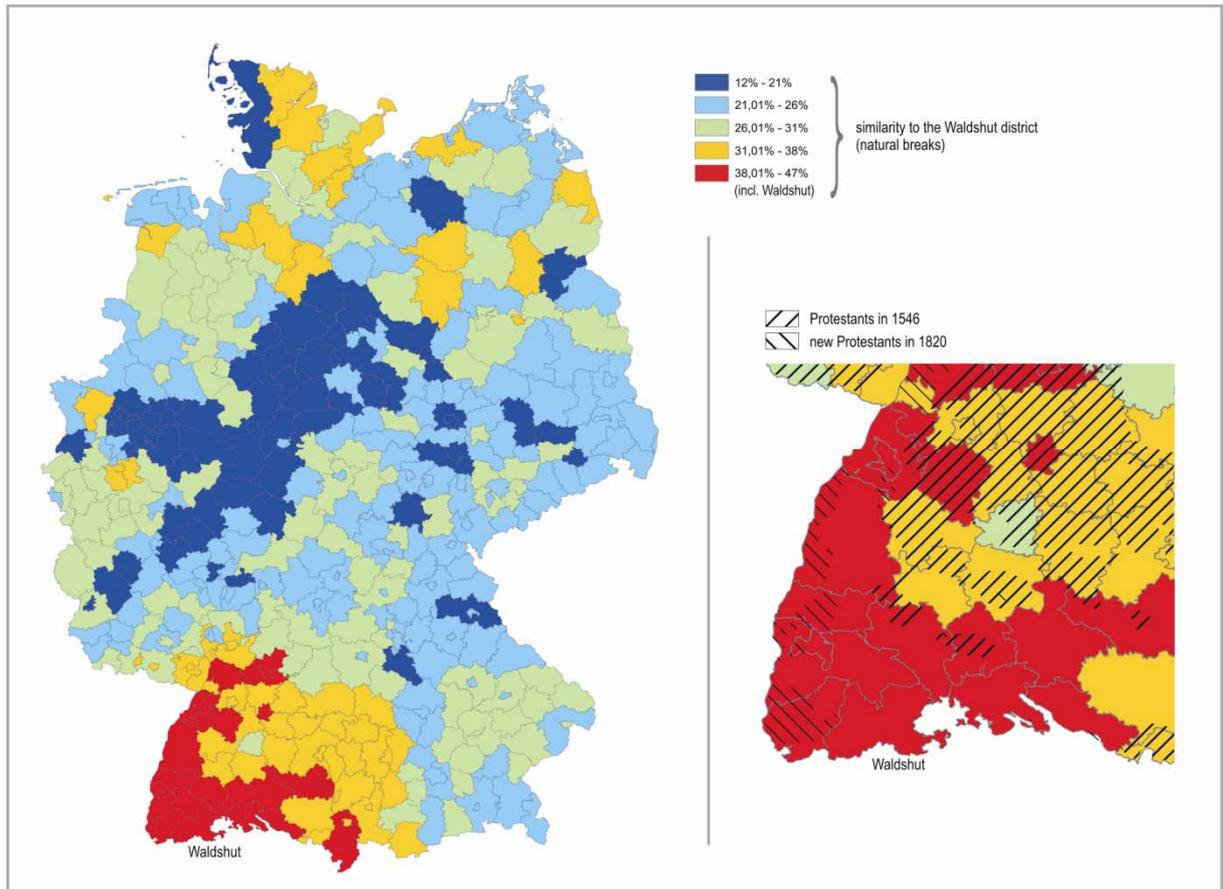
Notes: *Similarity of all districts to the reference point Waldshut (marked). Red indicates highest familiarity and yellow indicates higher familiarity, while the green and blue indicate less familiarity. Data on religious denomination are taken from Steger et al. (1989).*

Figure 3: The Language Enclave Goslar
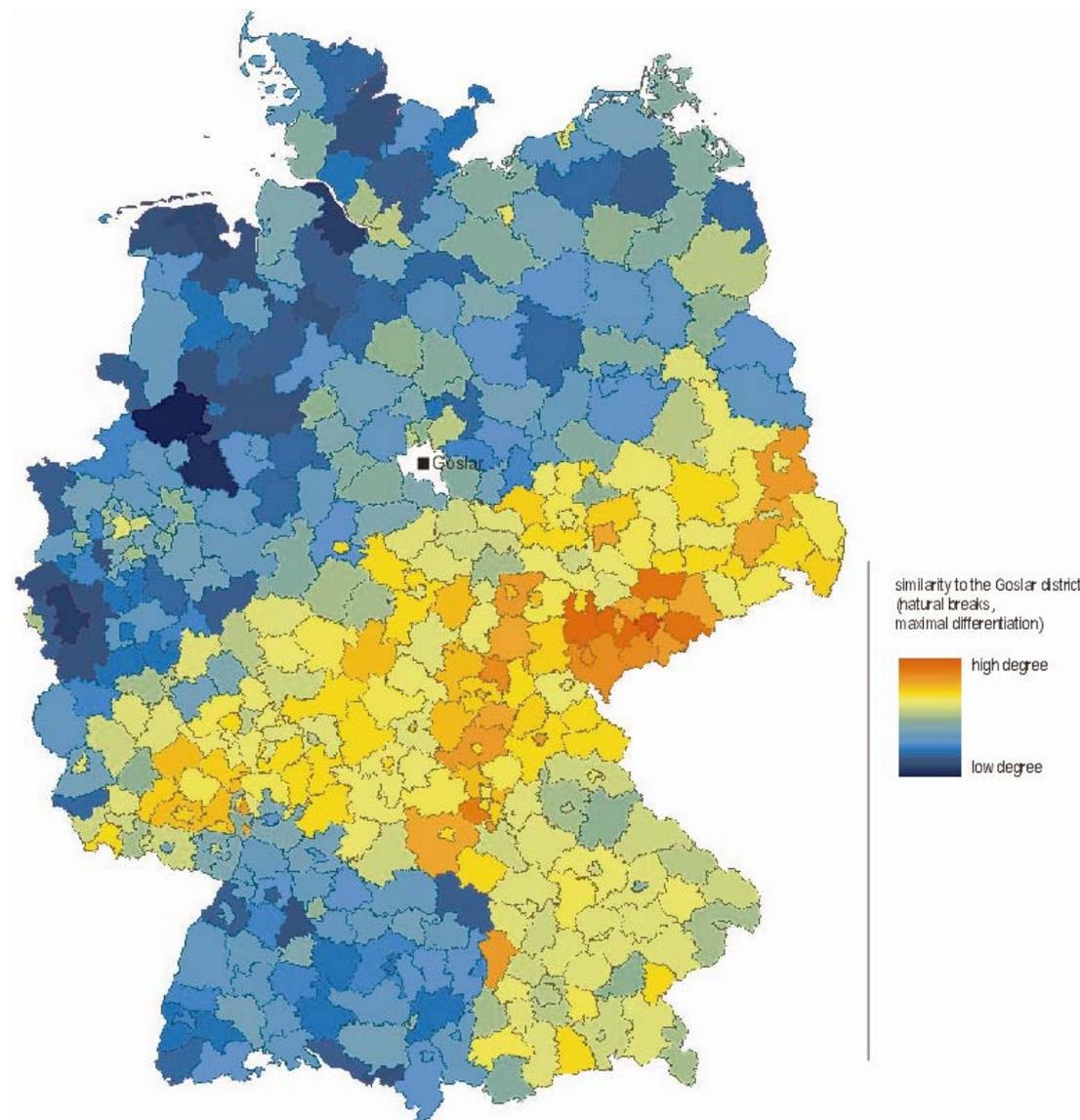
Notes: Similarity of all districts to the reference point Goslar (white spot). Red indicates highest familiarity and warmer tints (yellow and green) indicate higher familiarity, while the bluish tints indicate less familiarity.

Figure 4: Dialect similiarity with Passau – 19th century and today
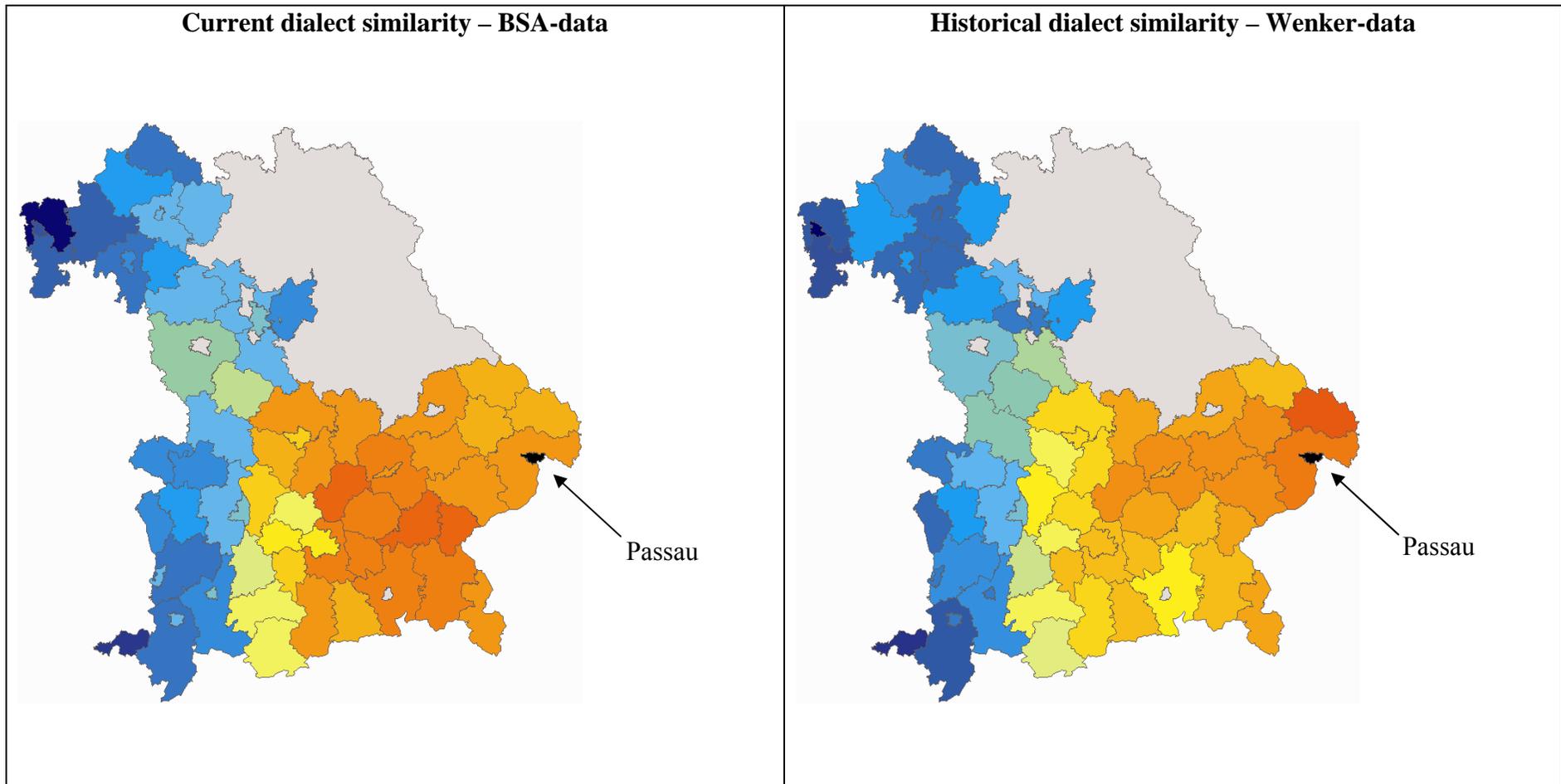
Notes: Similarity of all districts to the reference point Passau (black spot). Red indicates highest familiarity and warmer tints (yellow and green) indicate higher familiarity, while the bluish tints indicate less familiarity.

Table A1: Extended Data Description

Variable	Description and Source
Geographical Distance	The geographical distance between two districts is calculated as Euclidean distance in kilometers between each pair of districts' centroids.
Historical Political Borders	Historical borders refer to 38 member states and 4 independent cities that were part of the German Confederation at its foundation in 1815. Data are taken from a map in *Putzger – Historischer Weltatlas*, 89th edition, 1965. The dummy equals unity if a region pair does not belong to the same historic state, and zero otherwise.
Historical differences in industrial structure	Based on the German workplace and population census for the year 1907, we observe regional industry employment for 24 industry branches and for 97 functional regions (*Raumordnungsregionen, ROR*). Each of the 439 German districts can be uniquely assigned to one of those RORs. We calculate an dissimilarity index between all pairs of regions as follows: i) for each industry and ROR we calculate the share of that industry in total ROR employment, ii) for all industries and pairs of RORs we calculate the absolute value of the difference in the sectoral employment shares, iii) we sum those absolute differences across the 24 industries for each pair of RORs, iv) we then assign the resulting value of industrial dissimilarity to all districts within a ROR. That is, all districts within the same ROR have an index value equal to 0, while the index value is identical between all districts from some ROR_1 and all districts from some other ROR_2. Historical sectoral employment data are, unfortunately, not available on the geographical level of the 439 districts.
Historical travel distance	The measure is based on the route plan for a comprehensive German railway system outline by Friedrich List in 1833. Based on List's plans for a railway system, we define 12 major hubs (knots) and 22 spokes (railway lines) between or from these knots. We then surround this route system with a 50km corridor and determine each district as either being connected or not connected to the railway system, and we further determine which of the 22 railway line(s) intersect or border the district. Based on this categorization, we measure the historical travel distance as the number of knots between any pair of districts. District pairs in which at least one district is not intersecting or bordering any railway line are coded with the highest possible distance value of 7; all other district pairs have a distance value between 0 (at the same line) and 6. For district pairs in which at least one district is not intersecting or bordering more than one railway line, we choose the minimum distance.

Variable	Description and Source (continued)
Religious border dummy (1890)	The districts' historic shares of Catholics and Protestants in 1890 are calculated from a map in *Meyers Konversations Lexikon*, 4th edition, 1885–1892. The dummy equals unity if a region pair has different religious affiliations, i.e. an above average share of Catholics and Protestants respectively.

Soil	Soil concerns the main limitation to agricultural exploitation. The variable distinguishes between regions that have no limitation to agriculture and regions that have limitations due to less suitable soil characteristics.
	    1 no limitation to agricultural use
	    2 gravelly (over 35% gravels diameter < 7.5 cm)
	    3 stony (presence of stones diameter > 7.5 cm, impracticable mechanization)
	    4 lithic (coherent and hard rock within 50 cm)
	    5 concretionary (over 35% concretions diameter < 7.5 cm near the surface)
	    6 saline (electric conductivity > 4 mS.cm-1 within 100 cm)
	    7 others
	For our purpose, we collapse all limitations and create a binary variable that distinguishes regions that are more or less suitable for agriculture. The data stem from the European Soil Database (esdb) and are compiled by the European Soil Data Centre.

Slope	Slope is measured as the difference between the maximum and minimum elevations in meters. Flat regions are regions with a below average slope while steep regions are characterized by an above average slope.
Travel Distance (contemporaneous)	The travel distance is calculated in car minutes from one district's capital to the other.
Urban	This variable is based on a standard classification of German districts (*siedlungsstrukturelle Kreistypen*) according to their density and their spatial status (cf. Federal Office for Building and Regional Planning 2003). For our purpose, *urban areas* are districts characterized by a minimum city size of 100,000 inhabitants or a population density larger than 150 inhabitants per km². All other regions are classified as *peripheral areas*.