Why Are Educated and Risk-Loving Persons More Mobile Across Regions?

Stefan Bauernschuster[†], Oliver Falck[*], Stephan Heblich[‡], and Jens Suedekum[+]

Accepted at the Journal of Economic Behavior and Organization

Abstract

Why are better educated and more risk-friendly persons more mobile across regions? To answer this question, we use micro data on internal migrants from the German Socio-Economic Panel (SOEP) 2000–2006 and merge this information with a unique proxy for region-pair-specific cultural distances across German regions constructed from historical local dialect patterns. Our findings indicate that risk-loving and skilled people are more mobile over longer distances because they are more willing to cross cultural boundaries and move to regions that are culturally different from their homes. Other types of distance-related migration costs cannot explain the lower distance sensitivity of educated and risk-loving individuals.

[†] Department of Business Administration and Economics, University of Passau, Innstrasse 27, D-94032 Passau (Germany), Phone: +49 851-509-2540, Email: stefan.bauernschuster@uni-passau.de, Ifo Institute, and CESifo.

[*] Ifo Institute – Leibniz-Institute for Economic Research at the University of Munich, Poschingerstr. 5, D-81679 Munich (Germany), Phone: +49 89 9224 1370, Email: falck@ifo.de, University of Munich, and CESifo.

[‡] Division of Economics, University of Stirling, FK9 4LA, Stirling, UK, Phone: +44 1786 46 7481: Email: stephan.heblich@stir.ac.uk, IZA, and SERC (LSE).

[+] Mercator School of Management, University of Duisburg-Essen, Lotharstrasse 65, D-47057 Duisburg (Germany), Phone: +49 203 379 2357, Email: jens.suedekum@uni-due.de, CESifo, IZA, and SERC (LSE).

Acknowledgments: We thank William Neilson (the editor), an anonymous associate editor and referee, Joshua Gottlieb, Thierry Mayer, and seminar participants at the 1st European Meeting of the Urban Economics Association, the 6th Meeting of the Urban Economics Association, the NORFACE/CReAM Conference on "Migration: Economic Change, Social Challenge", and the 2012 Meeting of the German Economic Association for insightful comments and suggestions. The usual disclaimer applies.

1. Introduction

It is a well-established empirical fact that internal migrants—those who move across regions of the same country—move short distances significantly more than they move long distances. This finding of a detrimental effect of distance on regional migration dates back, at least, to the seminal studies of Sjaastad (1962) and Schwartz (1973) and has been confirmed for many different countries and time periods. It is also well known that highly educated individuals are more mobile in general, and also less sensitive to distance when they migrate, i.e., they move more easily to regions far from their homes.[1] Using survey data from the German Socio Economic Panel (SOEP), Jaeger *et al.* (2010) have recently shown that a similar point can be made for risk-loving persons who also tend to be more mobile across space. However, the reasons behind these mobility patterns are not yet well understood.

Two main hypotheses have emerged as explanation of these patterns. First, using Sjaastad's (1962) terminology, the adverse effect of distance on migration may result from *psychic* costs when leaving familiar surroundings. These are costs of having to adapt to a different regional culture (with different habits, norms, traditions, and so on), which tend to be higher for more distant destination regions. Second, individuals may be reluctant to move to distant regions because of direct "money costs" of migration, such as travel costs, or because they lack *information* about the prospective locations, along various dimensions such as the job and housing market, schools, facilities, and many other domains.

For both types of mobility costs, it can be argued that they affect individuals differently, depending on their level of education and their attitude towards risk. More educated and risk-friendly individuals may, for instance, be less sensitive to the *psychic* costs of migration because they can more easily adapt to (or are more willing to deal with) regional cultural differences. Similarly, better educated individuals may be more efficient in gathering information about prospective destination locations, while more risk-friendly persons may be more willing to encounter those various types of uncertainties.

A major and still unresolved problem in the literature on internal migration is that these hypotheses are difficult to disentangle. Both types of migration costs are distance-dependent, but neither of them is directly observable or measurable. It is therefore difficult to tear these explanations apart in order to understand *why* more educated and risk-friendly migrants overall move more easily over longer distances.

[1] A seminal paper on this issue is Dahl (2002). More recently, Malamud and Wozniak (2012) show that college education has a positive causal effect on interregional mobility in the United States, while Machin *et al*. (2012) establish a positive causal effect of the length of compulsory education on labor mobility in Norway. For Germany, Hunt (2004) shows that skilled migrants are more likely to move over longer distances.

In this paper we address this question by merging rich micro-data on internal migrants from the German SOEP with unique historical data on linguistic variation within Germany. These data stem from an encompassing language survey conducted by the linguist Georg Wenker between 1879 and 1888. They provide a unique opportunity to comprehensively measure cultural differences across German regions – something that would be very difficult, if not impossible, without linguistic data. In a gravity analysis, Falck *et al.* (2012) find that contemporaneous aggregate migration flows across German regions are lower—all else equal— the stronger the dialect difference between the origin and the destination region in the late 19th century. They then show that this represents the impact of intangible cultural barriers on regional migration in Germany.[2] However, Falck *et al.* (2012) only use aggregate migration flows in their study. We conduct our analysis at the micro level thus accounting for a host of individual characteristics of the (non-)movers.

Consistent with the previous literature, we first show that distance has a detrimental overall effect on migration. Furthermore, our analysis confirms that more educated and risk-loving individuals are more likely to migrate, and conditional on moving, they also tend to move over longer distances.[3] Our main contribution is that we shed light on the important question *why* this is the case.

The historical dialect data allow us to construct a *direct* (region-pair-specific) measure for cultural differences that are orthogonal to geographic distances, as well as a direct measure for *pure* geographic distances that are orthogonal to cultural differences. Put differently, we are able to derive a direct proxy for the cultural ("*psychic*") costs of migration, and a residual component that captures all other distance-dependent migration costs unrelated to culture. That latter, the *pure* geographic distances, thus encapsulate cross-regional travel and other direct migration costs, as well as the various types of information costs mentioned above. We then investigate to which concept of "distance" migrants are most sensitive.

Our main finding is that those *pure* geographic distances play no role in explaining the higher mobility of more educated and risk-loving persons. However, those individuals are systematically less sensitive to the cultural costs of migration. This lower sensitivity to cultural differences is thus the main explanation for the lower overall distance sensitivity in their migration decisions. To the best of our knowledge, ours is the first paper to provide

[2] Guiso *et al.* (2009) and Felbermayr and Toubal (2011) study the impact of cultural differences on cross-country trade and investment flows. The related approach by Falck *et al.* (2012) shows that cultural barriers to economic exchange also exist on a much finer geographically level, namely across regions of the same country.

[3] These results thus replicate the main findings of Jaeger *et al*. (2010), which is of interest in itself because we use more disaggregated data on internal migration in Germany than they do.

direct empirical evidence on the relative importance of these different costs of internal migration—an unresolved issue in the literature ever since Sjaastad (1962).

The rest of this paper is structured as follows. In Section 2 we describe our data. Section 3 presents the empirical approach and our baseline results. Section 4 is devoted to several robustness checks and extended analyses. Section 5 concludes.

2. Data

2.1. Contemporaneous migration data

We use data from the German Socio Economic Panel (SOEP), which is a large and representative household panel containing a rich set of socioeconomic variables (see Wagner et al. 2007). Specifically, we use a balanced panel of 10,393 individuals covering the period from 2000 until 2006. Of particular relevance for our purpose is the fact that individuals are followed not only over time but also across space. For every individual in the SOEP, we know the region of residence in the respective year, which allows us to recover regional migrations within Germany. Movers are identified as those who: (i) change their region of residence from one survey year to the next, and (ii) at the same time report having changed dwellings. We identify 994 individuals who moved at least once during the period of observation. Essentially, our SOEP data are comparable to the data used by Jaeger *et al.* (2010), but our analysis is conducted at a finer geographic level, i.e., at the level of the 439 German NUTS-3 regions (Landkreise), which are constructs roughly comparable to U.S. counties.

TABLE 1 HERE

We measure the movers' migration distances from the region of origin i to the destination region j. Our baseline measure is the simple linear distance (in km) between the geographical centers of the counties. We also have information about travel time by car (in minutes), which capture the regions' accessibility and are thus a good proxy for the actual travel costs between any pair of regions. On average, migrants moved 122 km (76 miles), which corresponds to a travel time of 114 minutes. Table 1 reports some further descriptive statistics for our sample of movers and non-movers, respectively. The table reveals patterns similar to those found by Jaeger *et al.* (2010). In particular, movers are on average younger, better educated, and also more risk-friendly than non-movers.[4]

[4] To measure individual risk aversion we use the risk indicator explained in detail in Jaeger *et al.* (2010). It is a dichotomous variable that takes the value of unity for individuals who rank their risk lovingness on a scale from 1 (very low) to 10 (very high) as 6 or higher.

2.2. Historical dialect data

Our main contribution in this paper is to separate the overall effect of distance on migration into two components: a cultural one ("psychic costs of migration") and a residual component that captures all geographic migration costs other than culture, such as travel costs or information costs for finding out about the destination's job and housing markets.

For this separation, we draw on a measure for region-pair-specific historical dialect similarity developed by Falck *et al.* (2012). That measure is based on unique linguistic data from a comprehensive language survey conducted by the linguist Georg Wenker between 1879 and 1888. The survey was intended to be an in-depth investigation of language variation within the newly created German Empire. At the time the survey was conducted, a standardized national language (*Hochdeutsch*) had not yet become prevalent; in fact, people even from neighboring villages sometimes were not able to properly communicate with each other. The survey asked pupils to read 40 German sentences, designed to reveal specific linguistic features, in their local dialect. In an extensive evaluation process, linguists have determined 66 prototypical characteristics that are most relevant for structuring the German language area. These characteristics have to do with the pronunciation of consonants and vowels as well as with grammar. These 66 characteristics are matched to Germany's current administrative classification scheme to quantify each region's dialect and to construct a dialect similarity matrix across all 439 regions.[5]

Figure 1 illustrates our approach. The map shows the regional similarities to the dialect spoken in *Marburg*, a region located roughly in the middle of Germany. The reference point *Marburg* is marked. Warm colors indicate a high, and cold colors a low, degree of linguistic similarity as measured in the late 19th century. By and large, it can be seen that regions closer to Marburg tended to have a more similar dialect than regions further away. However, the correlation between dialect distance and geographic distance is far from perfect. In particular, regions to the south and east tended to be linguistically much closer to Marburg's dialect than regions to the north and west. Stated differently, when drawing a circle around our reference point, it turns out that dialect distance to Marburg differs substantially across the geographically equidistant regions. The geography of dialects as recorded in the late 19th century thus apparently captures more than mere geographic distances, and our empirical approach exploits this variation offered by the linguistic data.

FIGURE 1 HERE

[5] See the Appendix for a more detailed explanation of the construction of the dialect similarity matrix.

What does dialect similarity capture? As is discussed at length in Falck *et al.* (2012), the geography of dialects reflects an entirety of historical interactions across the German regions from the centuries before. Influences such as common religious history, political borders, unique historical events, previous mass migration waves, etc., all left some long-lasting imprints on local dialects structures, and a higher degree of dialect similarity between any two regions indicates that those regions had more intensive interaction in the past, resulting in a higher degree of cultural similarity. There is, hence, a distance component in the dialect similarity measure: more adjacent regions tended to interact more in the course of history, and hence tended to develop more similar cultures and dialects. The dialect distances are, however, far from perfectly coincident with geographic distances, but provide a rich measure for the cultural similarity of German regions that would be very difficult, if not impossible, to capture without linguistic data.[6]

Today, dialects are far less common than they were in the 19th century when the language data were collected. Facilitated by linguistic diffusion, which is supported, e.g., by national media, individuals can now more easily communicate with each other in standard German, albeit with slightly different local accents. Nevertheless, even if dialects no longer create actual communication barriers, they are by far not nullified today but still reflect persistent *cultural* differences that have developed over centuries.[7] We therefore use the dialect differences from the late 19th century as our region-pair-specific measure for contemporaneous cultural differences.

The maximum number of linguistic correspondences that two regions can have is equal to 66 (see the data appendix). As is shown in Table 1, across all regional migrations that we have identified from the SOEP data, the average number of linguistic correspondences between the origin and the destination is 48. In other words, the average cultural cost that migrants encountered is 66 – 48 = 18, and the cultural cost of migration between two regions i and j is increasing in their historical dialect difference.

---

[6] Differences between national languages have often been used as a proxy for cultural differences across countries, see e.g. Ginsburgh and Weber (2011), Tabellini (2008) or Melitz (2008). The novel feature of Falck *et al.*'s (2012) and our study is that they analyze the variation of the *same* language across regions using detailed linguistic micro-data. To our knowledge, Grogger (2011) is the only other study which also exploits different speech patterns within the same language (English), but with a very different focus.

[7] The power of linguistic measures in revealing such deep cultural differences is widely discussed in other disciplines, including anthropology and sociology (see, e.g., Cavalli-Sforza 2000). Even if dialects are no longer actual barriers to communication in Germany, they continue to reflect the persistent cultural differences that developed in parallel to the language patterns over the long course of history. That is, differences in habits, norms, etc. are likely to be reflected in linguistic differences as well because those differences evolve in parallel with the process of cultural evolution. On this point, also see the recent contribution by Michalopoulos (2012).

2.3. Measuring cultural distance and pure geographical distance

We have shown that dialect differences across regions are correlated with, but capture more than geographic distance. To isolate the cultural component from the overall distances, we first regress the dialect similarity on the geographic distances across all pairs of regions *i* and *j*

$$dialect_{ij} = \alpha + \beta_1 distance_{ij} + \beta_2 traveltime_{ij} + \varepsilon_{ij} \tag{1a}$$

The results presented in Table 2a show that 41 percent of the variation in our dialect measure can be explained by geographic distance.[8] The residuals from this regression comprise all dialect differences that cannot be attributed to geographic distance. We hence take these residuals ε_{ij} (multiplied by -1) as our proxy for the pure cultural distances, which are by construction orthogonal to geographic distances.

TABLE 2 HERE

Analogously, we isolate the pure geographic distance component by regressing the measure of (linear) physical distance between regions *i* and *j* on their dialect similarity

$$distance_{ij} = \delta + \gamma_1 dialect_{ij} + \mu_{ij} \tag{1b}$$

The R^2 reported in Table 2b shows that 38 percent of the variation in geographic distance is coincident with our linguistic measure. The residuals from this regression (μ_{ij}) comprise the pure geographic distance purged of all cultural components.

Below we use the terms ε_{ij} and μ_{ij} as our baseline concepts of cultural distance and pure geographic distance, respectively, and investigate to which distance type migrants are more sensitive. To investigate whether our results are sensitive to the definition of the two distance concepts, we consider various extended and alternative specifications in the robustness checks below (see Section 4.1).

3. Empirical Analysis

3.1. The overall impact of distance on individual migration decisions

In a first step, we replicate the conventional approach of the existing literature and focus on the raw $distance_{ij}$ between origin and destination of the respective move. Specifically, we follow Jaeger *et al.* (2010) and model the decision to move as a dichotomous variable that

[8] We measure geographic distances both with linear physical distances and with travel time. Results would not change qualitatively if we captured geographic distance by only one of these concepts.

takes the value 1 for all individuals who moved from one region to another at least once in the period from 2000 until 2006; 0 otherwise. We then run simple probit regressions where we control for observable individual characteristics. The results, reported in Column (I) of Table 3, show that better educated, more risk loving, and younger individuals are more likely to move. Singles are also more likely to migrate. The main insights from the descriptive statistics (Table 1) are thus confirmed in this multivariate regression framework.[9] Both the willingness to take risk and education are important determinants of migration decisions, not only in statistical but also in economic terms. One more year of education, for example, raises the probability of moving by 0.9 percentage points. This is a substantial effect, given that just 9.6 percent of all individuals in our sample are movers. In terms of standard deviations, this means that a one standard deviation increase in years of education raises the probability of moving by roughly 2.3 percentage points.

We now focus on the subsample of movers. The results in column (II) of Table 3 show that, conditional on moving, better educated individuals move over longer (linear physical) distances. Note that all our distance-related outcome variables are z-standardized. One more year of education increases migration distance by 0.104 standard deviations. Put differently, a one standard deviation increase in years of education raises the migration distance by 0.29 standard deviations. The coefficient for risk lovingness is positive and large, yet imprecisely estimated. Interestingly, while we saw no effect of being from East Germany on the propensity to move (see Column (I)), we find that, conditional on moving, East Germans move greater distances. This can be explained by the fact that if East Germans move, they usually move to West Germany rather than within East Germany due to the large differences in per capita income and unemployment rates between East and West Germany that still exist today.[10] Finally, column (III) of Table 3 shows that the results are similar if we take z-standardized travel time (in minutes) as the outcome variable instead of physical distance.

TABLE 3 HERE

These findings are in agreement with the migration literature (see, e.g., Schwartz 1973), as well as with the recent findings by Jaeger *et al*. (2010) that more risk-loving persons are more mobile across space. Yet, it is unclear from Columns (II) and (III) of Table 3 *why* better educated and more risk-loving people are less distance sensitive in their migration decisions.

[9] All independent variables are measured for the year 2004 (the only year for which information on risk attitudes is available) while the period of observation for moving is 2000 until 2006. We explicitly choose this period of observation in order to be able to replicate the results of Jaeger *et al.* (2010).

[10] Indeed, of all individuals in our sample of movers who are born in East Germany, 45% have moved to West Germany by 2006, whereas only 4.5% of all individuals born in the West have moved to East Germany by 2006.

3.2. Main Results: Cultural versus pure geographic migration costs

The detrimental effect of distance on migration may be due to "psychic costs" capturing cultural differences across German regions, but it may also be due to other types of distance-related migration costs. To disentangle these different channels, we now use the two novel concepts of region-pair-specific distances—cultural distance (ε_{ij}) and pure geographic distance (μ_{ij}) — that we have constructed above.

The results are shown in Columns (IV) and (V) of Table 3, where our outcome variables are again z-standardized. Recall from Columns (II) and (III) that better educated and more risk-loving migrants are, overall, less sensitive to geographic distance. That is, conditional on moving, these individuals move more easily over longer distances. The results shown in Columns (IV) and (V) suggest that this overall effect is solely driven by the lower sensitivity of these individuals to cultural distance.

In Column (IV) the dependent variable is the cultural distance ε_{ij} between the origin and the destination region among all 994 migrants in the SOEP data. We find that better educated and more risk-loving individuals move more easily to destinations with a greater distance-adjusted dialect difference, i.e., to culturally less familiar environments. Both effects are highly statistically significant. In other words, these individuals are less sensitive to regional cultural differences than are lower skilled and more risk-averse persons.

Column (IV) also reveals a large and positive coefficient for the abroad dummy, that is, foreign-born individuals move on average to culturally more distant regions. This result is in line with our interpretation of historical dialect similarity being a proxy for persistent cultural similarity between German regions. Indeed, we would expect that the historical cultural imprints in a region are less relevant for the internal migration decision of foreign-born individuals than for native Germans.

In Column (V) the dependent variable is μ_{ij}, the pure geographic distance between the origin and the destination region purged of all cultural components. As can be seen, pure geographic distance seems to play a much weaker role. For risk-lovingness, the estimated coefficient is insignificant, for years of education it is barely significant and becomes unstable in robustness checks. The only clear result here is the previously mentioned one that East Germans, conditional on migrating, move over longer distances than West Germans, which remains true when distances are detached from the cultural component.

Summing up, the main reason why more educated and risk-loving persons are willing to move further away from their origin regions seems to be, that they are less sensitive to regional cultural differences. In other words, they seem to care less about the fact that other regions often have different traditions, habits, norms, and cultural backgrounds that they will have to deal with if they move to these destinations.[11] The higher mobility of skilled and risk-loving persons is, on the other hand, not well explained by the argument that they are less affected by other geographic migration costs, such as costs of information about the prospective destinations.

In the terminology of the traditional regional migration literature (Schwartz, 1973; Sjaastad 1962), our results therefore suggest that the "psychic costs" may actually be the most important type of migration costs, particularly for less educated and risk-averse individuals.[12]

4. Robustness Checks

4.1. Alternative specifications of the empirical analysis

We first address the robustness of our findings with respect to specification and estimation issues. A first concern is that the linear specification in equations (1a) and (1b) may be inappropriate, as it attributes any non-linearity in the spatial relationships to the error terms and, thereby, to our residual measures for cultural distance and pure geographic distance, respectively. To address this concern, we re-estimate (1a) and (1b) and include also quadratic terms for the different right-hand side distances in the estimations. It turns out that those non-linear terms are, indeed, statistically significant and that the R^2 levels increase to 0.467 and 0.384, respectively. We then use the residuals ε_{ij} and μ_{ij} from these regressions as our new distance measures, and re-estimate our baseline specification.

The results are reported in columns (I) and (II) of Table 4, which correspond to columns (IV) and (V) of Table 3 except for the different construction of cultural and pure geographic distance. As can be seen, even when taking those non-linear impacts into account, there is no change in our main conclusions: High-skilled and risk-loving persons are less sensitive to cultural distance, whereas pure geographic distance seems to play a lesser role.[13]

---

[11] We should emphasize once more that our proxy captures *cultural* differences across German regions, not linguistic or contemporaneous dialect differences per se, see footnote 7.

[12] Since all distance variables in Table 2 are z-standardized, we can also directly compare the magnitudes of the coefficients. As can be seen, the coefficients for years of education and risk lovingness are quite similar in Columns (II) and (IV); the difference in these coefficients is in fact statistically insignificant. This finding thus also suggests that the lower sensitivity to overall distances is driven by a lower sensitivity to cultural distance, whereas pure geographic distance is of second-order importance.

[13] Our findings also remain robust when using polynomials of even higher order in the construction of our residual measures ε_{ij} and μ_{ij}. Detailed results are available upon request from the authors.

Next, instead of using our two distance variables ε_{ij} and μ_{ij}, which are obtained as residuals from preceding regressions, we have also adopted a more direct approach. In particular, we use the raw physical migration distance of the internal moves as the outcome variable, and directly control for the dialect differences in the regression. Relatedly, we use the raw dialect distance as the dependent variable while controlling for the physical distance and travel time of the respective moves. Results are reported in columns (III)-(V) of Table 4.

The large and positive correlations between years of education and the risk indicator of the migrants, on the one hand, and the raw physical migration distance, on the other hand, completely disappear once we control for dialect differences (compare Columns (III) and (IV) of Table 4). However, as shown in Column (V), the positive and significant associations between dialect distance and years of schooling, as well as the risk indicator, remain robust even when we control for physical distance and travel time. Thus, high-skilled and risk-loving individuals are more likely to cross cultural borders even conditional on geographic distance, and this lower sensitivity to cultural differences appears to be the main reason why those individuals are more mobile across space overall.

TABLES 4 AND 5 HERE

Last, we conduct conditional logit estimations as an alternative empirical approach. Specifically, we build subsamples of high-skilled, low- and medium-skilled, risk-averse and risk-loving individuals, and then model the individual location decision as the choice between the different regions, while allowing the relevance of the characteristics of the potential destinations (our different distance measures) to differ across the subsamples.

For ease of computation, we aggregate our distance measures on the level of 97 planning regions (*Raumordnungsregionen*). As can be seen in Table 5, the results are fully consistent with our baseline findings. We find that raw geographic distance tends to be less relevant for high-skilled and risk-loving individuals than it is for low-skilled and risk-averse individuals. Turning to the reason for this pattern, we cannot reject the hypothesis that high-skilled and risk-loving individuals react similarly to pure geographic migration costs. However, we strongly reject the hypothesis that the same is true for the cultural costs of migration: rather, we find further evidence that high-skilled and risk-loving individuals are systematically less sensitive to cultural migration costs.

4.2. Economic differences between origin and the destination region

Returning to our benchmark specification as in Table 3 above, we now check if our results are confounded by region-specific economic differences between the origin and the destination, which may act as pull or push factors of individual migration decisions. Note that this would only be the case if these factors confound the education and risk coefficients systematically different across the regressions on cultural and pure geographic distance. To still address this issue, we include earnings per capita in the origin and destination and the pair-specific differences in the industrial structure, as migration flows may respond to those variables. Industry differences are derived from regional employment data from the German Social Insurance Statistics.[14] In unreported regressions (available upon request), we find that controlling for these additional variables does not affect our main results.

Our main result also remains robust when dropping all within-state movers and focusing on the subsample of individuals who moved from one Federal state to another. Although the number of observations drops considerably from 994 to 412 movers, our main results are unaffected. Vice versa, focusing only on within-state movers we also obtain results that are consistent with our baseline findings. This suggests that educated and risk-loving persons are less sensitive to cultural differences for both, long and short moves, and that our main results are not driven by just one type of internal migration flows.

4.3. Young migrants, endogenous origin locations and moves to big cities

The degree of regional labor mobility in Germany is considered to be relatively low, compared, e.g., to countries like the United States (Molloy *et al.*, 2011). That is, many Germans change regions only rarely (if at all) during their lifetimes, so that moves are typically regarded as major events for the respective individuals.[15] Unfortunately, we cannot observe the birth location of the SOEP respondents, or if the migrants are first-time movers. Despite the generally low degree of mobility, it may therefore be the case that that the observed origin location in the year 2000 is different from the region where the respective individual has his or her cultural roots, namely if he or she has moved within Germany before. If that were the case for a significant number of migrants in our data set, i.e., if the observed residence in 2000 was previously chosen for some unrelated reason (such as career concerns

[14] We generate a dissimilarity index between all pairs of regions that is calculated as the sum of the absolute differences between region *i* and region *j*'s employment shares across 59 different industries. Accordingly, larger values indicate stronger dissimilarity in regional industry structures.

[15] In our data set, less than 10% of the SOEP respondents were identified as movers over a time frame of six years (2000-2006).

or university choice), our dependent variables would be measured with error resulting in larger standard errors. However, this would only interfere with our empirical results if this measurement error was systematically different across our outcome variables.

To still investigate these issues, we focus on young individuals who are not older than 25 in 2000. These individuals are much more likely to be first-time movers who migrate away from their original place of birth. The results for this subsample are shown in Table 6a. As can be seen, the results for the subsample of young migrants are similar as the benchmark results from Table 3, although standard errors are larger because the number of observations drops considerably. Results are also similar to those obtained for the subsample of older migrants (aged 25 or above), which are shown in Table 6b. Our main results therefore seem to hold for migrants of different age groups. A related robustness check confirms that the results also hold for individuals without university education.

TABLES 6a, 6b AND 7 HERE

We have also taken into consideration that the observed moves of individuals may originate from locations that were chosen as temporary residences. More specifically, wherever possible, we have tried to recover the residences of the individuals at age 18 from pre-2000 waves of the SOEP. For the observed moves in the time window 2000-2006, we have then measured the migration distance not between the observed origin in 2000 and the ultimate destination region, but between the so constructed "birth location" (the residence observed at age 18) and the final destination. The results we obtain after conducting that exercise are reported in Table 7. They turn out to be similar to those reported in Table 6a. Summing up, even after taking into account that observed origin locations in the year 2000 may not be the cultural origin for every internal migrant, we obtain results in line with our baseline findings from Table 3.

TABLE 8 HERE

Finally, in a related robustness check, we investigate whether our results are driven by moves to big cities. Individuals might, for instance, temporarily move to the largest metropolitan areas, even if this does not match their cultural preferences, in order to benefit from better learning opportunities and career prospects that are typically much better there (see Glaeser and Maré 2001, Peri 2002). Among all migrants, 89 individuals in our sample moved to one of the five biggest German cities (Berlin, Hamburg, Munich, Cologne, and Frankfurt) during the period of observation. We drop those observations and re-run our regressions. As can be seen from Table 8, the results are again similar to those obtained with the full sample of movers.

5. Conclusions

In this paper, we have used unique historical data on local dialects to construct a *direct* (region-pair-specific) measure for cultural differences within Germany that is orthogonal to the conventional geographic distance measures, as well as a *direct* (region-pair-specific) measure for pure geographic distances within Germany that is orthogonal to cultural differences. Merging this information with the rich individual-level data from the German Socio-Economic Panel (SOEP), we have in a first step replicated Jaeger *et al.*'s (2010) finding that risk-loving and skilled migrants are more mobile over longer distances than risk-averse and low-skilled migrants. Extending that study and the extant literature on internal migration, we shed light on *why* this is the case. The main reason is that skilled and risk-loving persons are more willing to cross regional cultural boundaries and move to destinations that are culturally different from their homes. Pure geographic migration costs play only a minor role to explain this pattern across different types of individuals. These results are robust to a variety of specification tests and extended analyses. To the best of our knowledge, we are the first to provide direct empirical evidence on the relative importance of these different costs of internal migration. Our results suggest that more educated and risk-friendly individuals are less sensitive to the *psychic* costs of migration

Our paper contributes to a recent line of research showing that cultural differences matter for economic decisions even within a single country. Our findings show that cultural barriers between regions can impede internal migration particularly of less educated and risk-averse individuals. Thus, we find support for assumptions often made in the internal migration literature that previously could not be tested rigorously due to a lack of data capturing genuinely cultural dimensions.

Literature

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

Dahl, Gordon B. (2002), Mobility and the Return to Education: Testing a Roy Model with Multiple Markets, *Econometrica*, 70(6), 2367–2420

Falck, Oliver, Heblich, Stephan, Lameli, Alfred, and Jens Suedekum (2012), Dialects, Cultural Identity, and Economic Exchange, *Journal of Urban Economics*, 72(2-3), 225-239.

Felbermayr, Gabriel, and Farid Toubal (2010), Cultural Proximity and Trade, *European Economic Review*, 54(2), 279–293.

Ginsburgh, Victor and Shlomo Weber (2011), *How Many Languages Do We Need? The Economics of Linguistic Diversity*, Princeton: Princeton University Press.

Glaeser, Edward L., and David C. Maré (2001), Cities and Skills, *Journal of Labor Economics*, 19(2), 316-42.

Grogger, Jeffrey (2011), Speech Patterns and Racial Wage Inequality, *Journal of Human Resources*, 46(1), 1-25.

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

Hunt, Jennifer (2004), Are Migrants More Skilled than Non-Migrants? Repeat, Return, and Same-Employer Migrants, *Canadian Journal of Economics*, 37(4), 830–849.

Jaeger, David A., Dohmen, Thomas, Falk, Armin, Huffman, David, Sunde, Uwe, and Holger Bonin (2010), Direct Evidence on Risk-Attitudes and Migration, *Review of Economics and Statistics*, 92(3), 684–689.

Machin, Stephen, Pelkonen, Panu, and Kjell G. Salvanes (2012), Education and Mobility, *Journal of the European Economic Association,* 10 (2), 417-450.

Malamud, Ofer, and Abigail K. Wozniak (2012), The Impact of College on Migration: Evidence from the Vietnam Generation, *Journal of Human Resources*, 47(4), 913-950.

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

Michalopoulos, Stelios (2012), The Origins of Ethnolinguistic Diversity, *American Economic Review*, 102(4), 1508–39.

Molloy, Raven, Smith, Christopher, and Abigail Wozniak (2011), Internal Migration in the United States, *Journal of Economic Perspectives* 25(2), 1–42.

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

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

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

Tabellini, Guido (2008), Institutions and Culture, *Journal of the European Economic Association*, 6(2-3), 255–294.

Wagner, Gert, Frick, Joachim, and Jürgen Schupp (2007), The German Socio-Economic Panel Study (SOEP) - Scope, Evolution and Enhancements, *Schmollers Jahrbuch,* 127(1), 139–169.

Table 1: Descriptives

		All (I)	Non-movers (II)	Movers (III)
Physical distance in km	mean	11.68	0	122.17
	std.dev.	(58.99)	(0)	(151.35)
	N	10,393	9,399	994
Travel time in min	mean	10.93	0	114.33
	std.dev.	(49.27)	(0)	(116.50)
	N	10,393	9,399	994
Dialect similarity	mean	64.30	66	48.21
	std.dev.	(6.61)	(0)	(13.08)
	N	10,393	9,399	994
Cultural distance[a]	mean	16.10	17.14	6.30
	std.dev.	(4.10)	(0)	(8.32)
	N	10,393	9,399	994
Pure geographic distance[b]	mean	-12.20	-14.11	5.82
	std.dev.	(17.08)	(0)	(51.90)
	N	10,393	9,399	994
Risk index	mean	4.50	4.44	5.02
	std.dev.	(2.27)	(2.26)	(2.26)
	N	10,393	9,399	994
Age	mean	43.49	44.28	36.02
	std.dev.	(11.08)	(10.86)	(10.29)
	N	10,393	9,399	994
Years of education	mean	12.14	12.05	13.06
	std.dev.	(2.54)	(2.50)	(2.77)
	N	10,393	9,399	994
Place of origin				
West Germany	ratio in %	0.59	0.59	0.60
	N	6,137	5,538	599
East Germany	ratio in %	0.28	0.28	0.30
	N	2,923	2,623	300
Abroad	ratio in %	0.13	0.13	0.10
	N	1,333	1,238	95
Gender				
Female	ratio in %	0.52	0.52	0.52
	N	5,447	4,927	520
Male	ratio in %	0.48	0.48	0.48
	N	4,946	4,472	474
Marital status				
Married	ratio in %	0.70	0.73	0.49
	N	7,323	6,837	486
Single	ratio in %	0.30	0.27	0.51
	N	3,070	2,562	508

Notes: The table reports means, standard deviations and the number of observations, or ratios and the number of observations, of the respective variables. The sample consists of all individuals who took part in the SOEP surveys every year from 2000 to 2006 (Column (I)), on the subsample of all individuals who did not move between 2000 and 2006 (Columns (II)), and on the subsample of all individuals who moved between 2000 and 2006 (Column (III)). [a]"Cultural distance" refers to the residuals ε_{ij} from Equation (1a) and captures dialect distances purged of physical distances and travel times. [b]"Pure geographic distance" refers to the residuals μ_{ij} from Equation (1b) and captures physical distances purged of dialect distances.

Table 2a: Dialect similarity, geographic distance, and travel time

	Dialect similarity	
	coeff.	std.err.
Linear physical distance (in km)	0.014 ***	0.000
Travel time (in min)	-0.078 ***	0.001
Constant	48.86 ***	0.051
N	192,721	
R^2	0.412	

Notes: The table reports OLS coefficients and standard errors for a regression of dialect similarity on linear physical distance and travel time by car. The units of observation are region by region combinations. *** 1% level of significance, ** 5% level of significance, * 10% level of significance.

Table 2b: Linear physical distance and dialect similarity

	Linear physical distance (in km)	
	coeff.	std.err.
Dialect similarity	-8.847 ***	0.026
Constant	598.40 ***	0.886
N	192,721	
R^2	0.379	

Notes: The table reports OLS coefficients and standard errors for a regression of linear physical distance on dialect similarity. The units of observation are region by region combinations. *** 1% level of significance, ** 5% level of significance, * 10% level of significance.

Table 3: Determinants of moving and migration distances

	Move yes/no	Linear physical distance	Travel time	Cultural distance[a]	Pure geographic distance[b]
	Mfx	std.	std.	std.	std.
	(I)	(II)	(III)	(IV)	(V)
Years of education	**0.009** ***	**0.105** ***	**0.078** ***	**0.129** ***	**0.064***
	(0.001)	(0.029)	(0.027)	(0.023)	(0.035)
Risk indicator	**0.023** ***	**0.222**	**0.202**	**0.283** **	**0.009**
	(0.006)	(0.165)	(0.152)	(0.131)	(0.198)
Age	-0.004 ***	0.016 *	0.018 **	-0.007	0.015
	(0.000)	(0.008)	(0.008)	(0.007)	(0.010)
Female	0.003	0.169	0.178	0.127	0.025
	(0.005)	(0.163)	(0.150)	(0.129)	(0.208)
Married	-0.041 ***	0.010	0.048	-0.204	0.025
	(0.007)	(0.174)	(0.160)	(0.137)	(0.208)
Place of origin (omitted category: West Germany)					
East Germany	-0.000	0.852 ***	0.871 ***	-0.036	0.716***
	(0.006)	(0.179)	(0.165)	(0.142)	(0.215)
Abroad	-0.004	-0.193	-0.226	0.433 *	-0.290
	(0.008)	(0.284)	(0.261)	(0.224)	(0.341)
N	10,393	994	994	994	994
Log likelihood	-2,935				
R^2		0.043	0.047	0.045	0.019

Notes: The table reports marginal effects of probit regressions evaluated at the sample mean (Column (I)) and OLS coefficients (Columns (II) through (V)). The estimations are run on the sample of all individuals who took part in the SOEP surveys every year from 2000 to 2006 (Column (I)) and on the subsample of all individuals who moved between 2000 and 2006 (Columns (II) through (V)). Outcome variables in Columns (II) through (V) are standardized to a mean of 0 and a standard deviation of 1. All outcome variables are coded such that higher values signify greater distance. The two main variables of interest are years of education and a risk indicator taking the value of unity for individuals who rank their risk lovingness on a scale from 1 (very low) to 10 (very high) as 6 or higher. Standard errors are given in parentheses; *** 1% level of significance, ** 5% level of significance, * 10% level of significance.

[a]"Cultural distance" refers to the recoded and standardized residuals ε_{ij} from Equation (1a) and captures dialect distances purged of physical distances and travel times. [b]"Pure geographic distance" refers to the standardized residuals μ_{ij} from Equation (1b) and captures physical distances purged of dialect distances.

Table 4: Alternative specification including distance controls

	Cultural distance std. (I)	Pure geographic distance std. (II)	Linear physical distance std. (III)	Linear physical distance std. (IV)	Dialect distance std. (V)
Years of education	**0.119*****	**0.055**	**0.105 *****	**- 0.010**	**0.054 *****
	(0.025)	(0.036)	(0.029)	(0.017)	(0.013)
Risk indicator	**0.249***	**0.036**	**0.222**	**- 0.053**	**0.121 ***
	(0.140)	(0.203)	(0.165)	(0.097)	(0.072)
Other controls	YES	YES	YES	YES	YES
Dialect distance std.				**1.049*****	
				(0.024)	
Linear physical distance std.					0.060
					(0.071)
Travel time std.					**0.630*****
					(0.077)
N	994	994	994	994	994
Log likelihood					
R^2	0.049	0.011	0.043	0.673	0.692

Notes: The table reports OLS coefficients. The estimations are run on the subsample of all individuals who took part in the SOEP surveys every year from 2000 to 2006 and moved between 2000 and 2006. Outcome variables in all columns are standardized to a mean of 0 and a standard deviation of 1. All outcome variables are coded such that higher values signify greater distance. The two main variables of interest are years of education and a risk indicator taking the value of unity for individuals who rank their risk lovingness on a scale from 1 (very low) to 10 (very high) as 6 or higher. Columns (I) and (II) correspond to columns (IV) and (V) from Table 3 and use residuals from versions of the preceding regressions (1a) and (1b) which include linear and quadratic terms of distance in (1a) and dialect distance in (1b). Column (III) corresponds to Column (II) of Table 3. In Column (IV) the outcome variable is the raw linear physical distance associated with the move and we control for dialect distance. Analogously, the outcome variable in Column (V) is the raw dialect distance and we control for for linear physical distance and travel time. Standard errors are given in parentheses; *** 1% level of significance, ** 5% level of significance, * 10% level of significance.

Table 5: Conditional logit estimations on subsamples

	Conditional logit coeff.	std.err.		Conditional logit coeff.	std.err.
Linear physical distance			**Cultural distance** [a)]		
Risk-averse	-8.879***	0.199	Risk-averse	3.880***	0.043
Risk-friendly	-7.294***	0.210	Risk-friendly	3.517***	0.061
Test for equality of coefficients			*Test for equality of coefficients*		
chi²	30.10		*chi²*	23.74	
Prob>chi²	0.000		*Prob>chi²*	0.000	
Low- and medium-skilled	-9.307***	0.210	Low- and medium-skilled	3.974***	0.041
High-skilled	-6.676***	0.199	High-skilled	3.320***	0.063
Test for equality of coefficients			*Test for equality of coefficients*		
chi²	82.96		*chi²*	74.85	
Prob>chi²	0.000		*Prob>chi²*	0.000	
Travel time			**Pure geographic distance** [b)]		
Risk-averse	-6.141***	0.107	Risk-averse	0.197***	0.007
Risk-friendly	-5.248***	0.113	Risk-friendly	0.202***	0.011
Test for equality of coefficients			*Test for equality of coefficients*		
chi²	32.74		*chi²*	0.16	
Prob>chi²	0.000		*Prob>chi²*	0.690	
Low- and medium-skilled	-6.357***	0.113	Low- and medium-skilled	0.202***	0.007
High-skilled	-4.931***	0.108	High-skilled	0.191***	0.011
Test for equality of coefficients			*Test for equality of coefficients*		
chi²	83.21		*chi²*	0.57	
Prob>chi²	0.000		*Prob>chi²*	0.452	

Notes: The table reports conditional logit coefficients and standard errors. The estimations are run on four subsamples (risk-averse, risk-friendly, low- and medium-skilled, high-skilled) of the balanced panel of individuals who took part in the SOEP surveys from 2000 to 2006. High-skilled individuals are individuals with more than 13 years of schooling. Risk-friendly individuals are individuals who rank their risk lovingness on a scale from 1 (very low) to 10 (very high) as 6 or higher. The main variables of interest (geographic distance, travel time, cultural distance, pure geographic distance) are standardized to a mean of 0 and a standard deviation of 1. *** 1% level of significance, ** 5% level of significance, * 10% level of significance.
[a], [b]: See notes to Table 3.

Table 6a: Determinants of moving and migration distances: sample <=25 years old

	Move yes/no Mfx (I)	Linear physical distance std. (II)	Travel time std. (III)	Cultural distance[a] std. (IV)	Pure geographic distance[b] std. (V)
Years of education	**0.024 ***** (0.006)	**0.131 **** (0.060)	**0.115 **** (0.055)	**0.149 ***** (0.054)	**0.066** (0.069)
Risk indicator	**0.036** (0.026)	**0.025** (0.272)	**0.055** (0.251)	**0.229** (0.247)	**-0.234** (0.313)
N	1,237	314	314	314	314
Log likelihood	-685.9				
R^2		0.057	0.070	0.035	0.027

Table 6b: Determinants of moving and migration distances: sample >=25 years old

	Move yes/no mfx (I)	Linear physical distance std. (II)	Travel time std. (III)	Cultural distance[a] std. (IV)	Pure geographic distance[b] std. (V)
Years of education	**0.008 ***** (0.001)	**0.109 ***** (0.034)	**0.082 ***** (0.032)	**0.132 ***** (0.026)	**0.054** (0.042)
Risk indicator	**0.022 ***** (0.006)	**0.227** (0.202)	**0.196** (0.185)	**0.272 *** (0.152)	**0.060** (0.247)
N	9,328	719	719	719	719
Log likelihood	-2,324				
R^2		0.044	0.039	0.054	0.021

Notes: The specifications reported in panels A and B are analogous to those reported in Table 3, see the notes there for a more detailed description. The estimations in Panel A are run on the sample of all individuals not older than 25 in 2000 who took part in the SOEP surveys every year from 2000 to 2006 (Column (I)) and on the subsample of all individuals who were not older than 25 in 2000 and moved between 2000 and 2006 (Columns (II) through (V)). The estimations in Panel B are run on the sample of all individuals who took part in the SOEP surveys every year from 2000 to 2006 and were at least 25 years old in 2000 (Column (I)) and on the subsample of all individuals who were at least 25 years old in 2000 and moved between 2000 and 2006 (Columns (II) through (V)).

[a], [b]: See notes to Table 3.

Table 7: Determinants of moving and migration distances: origin region is region where individual lived at the age of 18

	Linear physical distance std. (I)	Travel time std. (II)	Cultural distance[a] std. (III)	Pure geographic distance[b] std. (IV)
Years of education	**0.078 ****	**0.073 ****	**0.110 *****	**0.008**
	(0.038)	(0.036)	(0.032)	(0.042)
Risk indicator	**-0.157**	**-0.147**	**0.001**	**-0.062**
	(0.185)	(0.172)	(0.155)	(0.203)
N	256	256	256	256
Log likelihood				
R^2	0.072	0.081	0.067	0.015

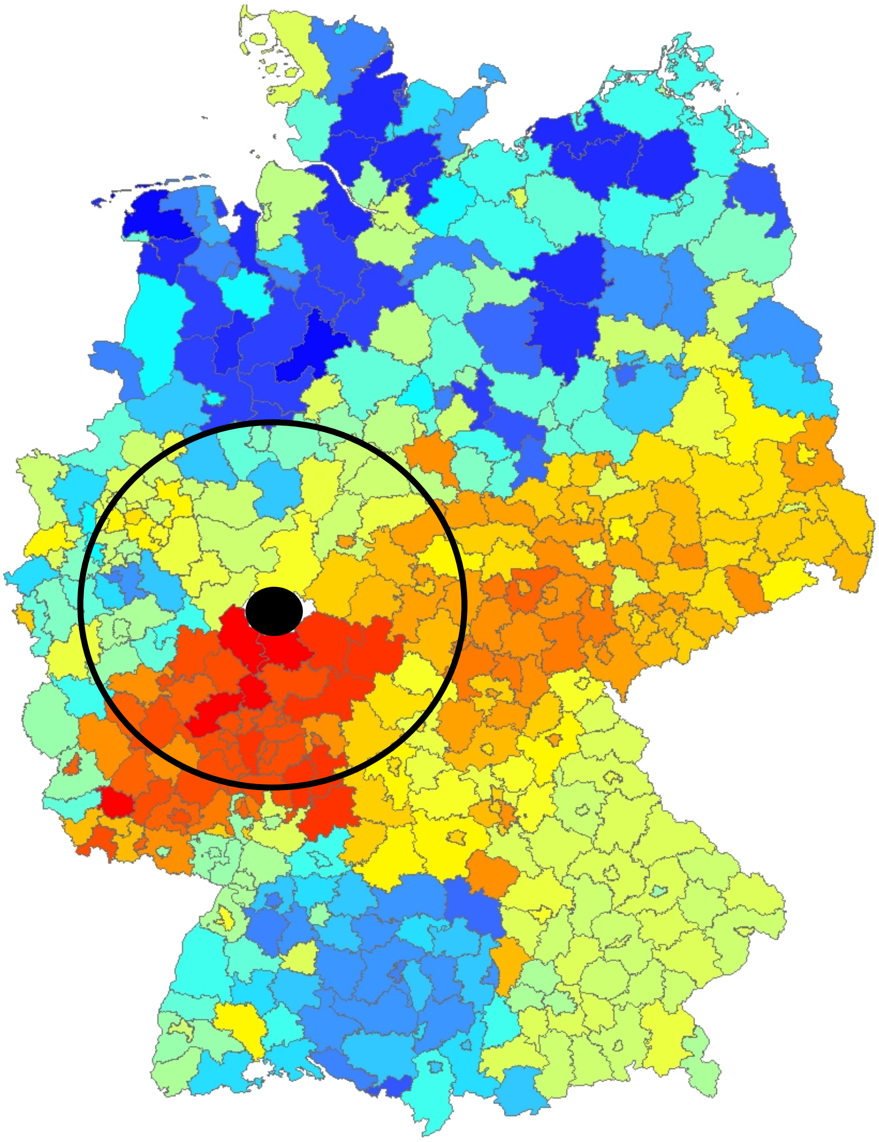
Notes: The specifications reported in this table are analogous to columns (II)-(V) in Table 3, see the notes there for a more detailed description. For the subsample of all individuals who moved between 2000 and 2006, we construct the migration distance not between the origin and destination of the move, but between the "birth place" (observed residence at age 18 according to previous SOEP-waves) and the final destination of the respective individual.
[a],[b]: See notes to Table 3.

Table 8: Determinants of moving and migration distances (without moves to Berlin, Hamburg, Munich, Cologne, or Frankfurt)

	Move yes/no mfx (I)	Linear physical distance std. (II)	Travel time std. (III)	Cultural distance[a] std. (IV)	Pure geographic distance[b] std. (V)
Years of education	**0.007 *****	**0.072 ****	**0.054 ****	**0.118 *****	**0.037**
	(0.001)	(0.029)	(0.027)	(0.025)	(0.033)
Risk indicator	**0.022 *****	**0.138**	**0.121**	**0.251 ***	**-0.123**
	(0.006)	(0.162)	(0.151)	(0.139)	(0.182)
N	10,304	905	905	905	905
Log likelihood	-2,777				
R^2		0.051	0.057	0.040	0.030

Notes: The specifications reported in this table are analogous to those reported in Table 3, see the notes there for a more detailed description. The estimations are run on the sample of all individuals who took part in the SOEP surveys every year from 2000 to 2006 (Column (I)) and on the subsample of all individuals who moved between 2000 and 2006 (Columns (II) through (V)); we drop all individuals who moved to one of the five big German cities of Berlin, Hamburg, Munich, Cologne, and Frankfurt.
[a],[b]: See notes to Table 3.

Figure 1: Dialect similarity — the case of Marburg



Notes: The figure shows dialect similarity of all districts to the reference point Marburg (marked). Degrees of dialect similarity (from highest to lowest) are indicated by: red, yellow, green, blue.

Data Appendix

Falck *et al.*'s (2012) dialect similarity matrix is constructed from 66 prototypical characteristics that are most relevant for structuring the German language area. These characteristics have to do with the pronunciation of consonants and vowels as well as with grammar, such as the use of accusative or dative in certain circumstances. For each district, the specific dialect is identified in the form of binary variables. One prototypical characteristic is the German word for *pound*. In the eastern parts of Germany, pound is mostly pronounced as "Fund," in the northern areas as "Pund," and in the southern parts as "Pfund." This leads to the following binary code: "Fund" = {1 0 0}; "Pund" = {0 1 0}; "Pfund' = {0 0 1}. To assign one binary code per language characteristic to each of the 439 regions in Germany, the individual linguistic map for the language characteristic—in the example the word *pound*—is layered over an administrative map of German regions by means of GIS software. This procedure is unambiguous if the entire region is characterized by the same pronunciation, which typically is the case. If more than one particular language characteristic is observed within a region, the most frequent variant is considered to be representative. Linguistic plausibility tests and cross-checks with the underlying raw data additionally assure the quality of these matches.

This procedure is repeated for all 66 language characteristics. The binary codes of the characteristics have between 2 and 18 realizations. This results in K=383 binary variables representing the dialect that was spoken in the area of a region in the late 19th century. Formally, the historical dialect of a current region *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 element of the vector is a binary variable [0,1]. Using this information, a dialect similarity matrix can be constructed across all *R*=439 regions. For 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, their similarity is quantified as overlap of these two vectors by means of a simple count similarity measure, $\ell_{rs} = \mathbf{i}^r \times \mathbf{i}^s$, where $0 \leq \ell_{rs} \leq 66$ for $r \neq s$ and $\ell_{rr} = 66$. The resulting matrix across all regions has dimension 439×439 with generic elements ℓ_{rs}.