Ruhose, Jens; Falck, Oliver; Lameli, Alfred

Conference Paper

The Cost of Migrating to a Culturally Different Location

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Employment Contracts, No. C04-V2

Provided in Cooperation with:
Verein für Socialpolitik / German Economic Association

Suggested Citation: Ruhose, Jens; Falck, Oliver; Lameli, Alfred (2014) : The Cost of Migrating to a Culturally Different Location, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Employment Contracts, No. C04-V2, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:
https://hdl.handle.net/10419/100327

The Cost of Migrating to a Culturally Different Location

Oliver Falck[†], Alfred Lameli[+] and Jens Ruhose[*]

This version: February 2014

Abstract

Ever since Sjaastad (1962), researchers have struggled to quantify the psychic cost of migration. We monetize psychic cost as the wage premium for moving to a culturally different location. We combine administrative social security panel data with a proxy for cultural difference based on unique data on historical dialect dissimilarity between German counties. Conditional on geographic distance and pre-migration wage profiles, we find that migrants demand a wage premium of about 1 percent for overcoming one standard deviation in cultural dissimilarity. The effect is driven by males, more pronounced for geographically short moves, and persistent over time.

Keywords: Migration Costs, Culture, Internal Migration, Psychic Cost
JEL codes: D51, J61, R23

[†] 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.

[+] Research Centre Deutscher Sprachatlas, Hermann-Jacobsohn-Weg 3, D-35032 Marburg (Germany), Phone: +49 6421 28 22482, Email: lameli@uni-marburg.de.

[*] Ifo Institute – Leibniz-Institute for Economic Research at the University of Munich, Poschingerstr. 5, D-81679 Munich (Germany), Phone: +49 89 9224 1388, Email: ruhose@ifo.de.

Acknowledgments: We thank Constantin Mang, Fabian Waldinger, Simon Wiederhold, Ludger Woessmann, and seminar participants at the Ifo Institute, the 3rd workshop for regional economics in Dresden, and the University of Luxembourg for helpful comments and discussion.

1. Introduction

The decision to migrate, and where exactly to go, is determined by comparing the costs and benefits of moving to the costs and benefits of alternatives. Benefits and costs can be monetary or non-monetary (Sjastaad 1962). Non-monetary migration costs include the psychic cost of moving from a familiar to a strange surrounding. As pointed out by Sjaastad, psychic cost is a result of tastes, which should be taken as given. Consequently, migration incentive transfers to compensate for this psychic cost would be inappropriate. Nevertheless, it is important to quantify psychic cost when analyzing rates of return to migration; otherwise, the rate of return to resources allocated to migration is biased.

Geographic distance between the place of origin and destination is viewed as a catch-all proxy for various costs of migration, including psychic cost (cf. Greenwood 1975). Schwartz (1973, p. 1161) argues that "psychic cost can be transformed into permanent transportation cost by figuring the needed frequency of visits to the place of origin so as to negate the agony of departure from family and friends." He further argues that this frequency increases with age but not with education. As a consequence, the interaction of distance with age indirectly measures the importance of the psychic cost of migration. Dahl and Sorenson (2010) study internal migration of skilled technical workers in Denmark and consider various distance measures, i.e., distance to parents and siblings, distance to the place where the individual grew up, and distance to prior residence. The relative importance of the various distance measures is an indication of the relative importance of psychic cost as compared to transportation cost. A non-distance-related approach that sheds light on the importance of the psychic cost of migration is a study by Barrett and Mosca (2013) finding that Irish male return migrants are more likely to suffer from alcohol problems than those who never moved.

We monetize the psychic cost of migration as the wage premium that migrants demand when moving to a culturally different location. For this purpose, we use administrative social security panel data in order to identify internal migrants in Germany. Internal migrants are defined as job switchers for whom the job switch involves moving house from one county to another, i.e., we do not study commuting. We merge the internal migrants' wage profiles over time with information on the geographic and cultural distance between their counties of origin and destination. Cultural distance is calculated from unique data on historical dialect dissimilarity between German counties (Falck *et al.* 2012).

This historical dialect information comes from a government-funded dialect survey conducted in the German Empire at the end of the 19th century. At this time dialects were still the prevalent medium for oral communication, often leading to significant problems in mutual intelligibility for people from different regions or even nearby towns of the German language area. As the most prominent expression of social identity, almost like a genome, historical dialects stored information about past interactions across German counties over time. Our broad and evolutionary perspective of culture is thus similar to that of Guiso *et al.* (2006),

who define culture as "those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation." The linguistic situation changed when social and economic exchange was intensified after the founding of the empire. Then the national language (*Hochdeutsch*) that as yet was mostly reserved to written contexts became more and more accepted. At the same time, and considerably increasing after World War II, German dialects show both symptoms of convergence and linguistic transfer from the national language. To get most explicit cultural consolidations at a very small geographic scale thus benefits from the use of historical dialect data.

Our findings imply that, conditional on geographic distance and pre-migration quarterly wage profiles, internal migrants demand a wage premium of about 1 percent for overcoming one standard deviation in dialect distance. This wage premium is most likely a lower-bound estimate for internal migrants since the county of an immediate origin of an internal migrant is not necessarily the place where he or she was born and socialized. For those cases, however, we would not expect to find a systematic correlation between wage changes and dialect distance. The effect is driven by males, more pronounced for geographically short moves, and persistent over time. Considering higher polynomial functions of geographic distance in the regressions provides additional confidence that the effect of dialect distance does not reflect only a non-linearity in the geographic distance effect. We also analyze those who have made multiple moves within a relatively short period of time and find that internal migrants who made a "wrong decision" in the first move correct this decision in the second move and demand a much higher wage premium.

The remainder of the paper is organized as follows. Section 2 introduces the wage data for internal migrants in Germany. Section 3 describes the historical dialect data. Section 4 explains our estimation strategy. Section 5 shows the baseline results. Section 6 provides a series of robustness checks. Section 7 reveals important effect heterogeneities with respect to gender, education, geographical distance, and multiple times movers. Section 8 investigates the long-term impact of cultural distance. Section 9 concludes.

2. Internal Migration in Germany

Information on internal migration in Germany stems from the IAB Employment Panel. Based on the quarterly employment statistics of the Federal Employment Agency, the panel covers a subsample of 2 percent from the universe of employees who are subject to social security in Germany. Besides gross monthly wages, the data provide background characteristics on age, gender, educational attainment, nationality, and place of work and residence.[1] Our sample period covers the years 1998 to 2006 and thus includes about 26 million quarterly observations from around 925,000 individuals. Since information on hours worked is not accurate in the IAB Employment Panel, we restrict our analysis to full-time employed

[1] To obtain the regional identifiers for the county of work and county of residence, we use the confidential weakly anonymous version of the scientific use file (see Schmucker and Seth 2009).

individuals. However, there are still workers who receive zero wages even if they are full-time employed. We follow Card *et al.* (2013) and drop all workers with daily wages below 10 euro. Another problem is that the wage data is top-coded at the social security maximum. The number of affected workers in the full sample is at the order of 10 to 12 percent of male workers and 1 to 3 percent for female workers (Card *et al.* (2013)). The literature has proposed to impute the missing wage information by assuming a normal wage distribution (Dustmann *et al.* (2009), Card *et al.* (2013)). However, we restrict our sample only to persons who have moved between two quarters. We find that only about 2 percent, either before or after the move, are top-coded. Thus, in total, we have only about 4 percent of top-coded observations. Therefore, instead of using imputation methods, we decided to check the robustness of our results by excluding this group and find that the results do not change.[2] Finally, we restrict our sample to German citizens only.

We define internal migrants as individuals who have changed their county of residence and their county of work between two consecutive quarters. In some cases, the information on county of residence and work is missing. In these cases, we allow for an administrative lag of one quarter and determine whether the person has moved by comparing the work and residence status of the person in the wave before the missing entry with the wave after the missing entry.[3] Our final sample contains 9,090 internal migrants. The internal migration rate in our sample is roughly 3 percent, which is comparable to official aggregate internal migration statistics for Germany.[4]

Panel A of Table 1 shows the distribution of wages four quarters before (*t-1, t-2, t-3, t-4)* and one quarter after the move (*t+1*). The average monthly nominal gross wage before the move is € 2,549 and increases by 4.3 percent to € 2,659 after the move.

<< Table 1 about here >>

However, the nominal wage differences across regions are probably not the relevant unit at which people look when they decide where to move. Local amenities, like good schools, transport infrastructure, health care providers, shopping alternatives, or leisure facilities are valued as well. In addition, the general price level for goods and services is important to determine the purchasing power of the wage.

[2] The social security data should only record wages until the social security maximum. However, there are a few cases in which the reported wage exceeded this number. We restricted these cases back to the social security maximum. We also performed robustness checks by omitting the bottom and top 5 percent of the wage distribution and the results are not sensitive.

[3] Omitting individuals with an administrative lag from the sample or controlling for them with an indicator variable does not change the results.

[4] The average overall internal migration rate for the period 1998 to 2006 was 4.6 percent (own calculations based on official migration and population data of the Federal Statistical Office 2013). Since our sample consists of working individuals subject to social security only, the internal migration rate in our sample is slightly lower.

Rental rates are an important component of the general price level in Germany. Almost 1/3 of the consumer price index is due to rental rates. Moreover, local amenities capitalize through demand factors into rental rates. Cities like Munich have a plenty of local amenities but also a high overall price level which feeds into the highest local rents (10.74 euros/sqm). [5] Whereas cities like Hof does not provide as many amenities and have a rather low consumer price index and therefore requires much lower rents (3.77 euros/sqm).

Correcting for the price level, we want to have the wage at the same purchasing power between the two cities. But we also want to compensate for the differences in amenities. Therefore, we inflate the wages in Hof to the wages in Munich by the relative rental price between Hof and Munich, 3.77/10.74 = 0.35. This procedure should equalize the different price and amenity levels between the cities such that people from Hof could afford the same goods and services and amenities as in Munich.

The average monthly indexed gross wage before the move is 4,623 and increases by 1.7 percent to 4,703 after the move (see Panel A of Table 1). The resulting measure of the indexed wage is, however, difficult to interpret because suitable comparisons are not available. To get a sense of the effect size and the robustness of our results, we will estimate the model with log wages and control for the rental prices on the right-hand side of the regression.

Another observation from Table 1, Panel A is that the average wage for movers to counties above the median dialect distance is higher than for movers to counties below the median dialect distance. This is the case not only one quarter after the move but at each point in time. This suggests that these people have higher skills than movers who do not move far away in cultural terms. But these skills also show up in wages prior to the move.

Panel B of Table 1 shows descriptive statistics for the distance data. The geographic distance certainly correlates with dialect distance. The mean geographic distance is 318 km for individuals who have moved to a county above the median dialect distance whereas the destination county is only 76 km away for the below-median mover.

This selection of individuals into moving across cultural borders can also be seen in Panel C of Table 1. Above-median movers are much more likely to have university degree (33.9 percent vs. 26 percent) and to have attended the highest academic track in secondary school (15.2 percent vs. 13.5 percent). There is also a gender gap. The share of male migrants is higher in the below-median mover group as in the above-median mover group. Age and (potential) experience are comparable across both groups.[6] The average age of the movers is

[5] We use rental prices averaged over the years 2004 to 2008 reported by the Federal Institute for Research on Building, Urban Affairs and Spatial Development, as well as the IDN ImmoDaten GmbH. The rental prices are turned into a price index expressed in terms of the most expensive place, Munich. Munich receives a value of one and all other counties are ranked relatively to Munich.

[6] Potential labor market experience is computed by *Age – 6 – Years of Education.*

32. The question is whether the wage around 30 is a meaningful measure of the productivity or life-time earnings of the individual. Studies by Haider and Solon (2006) who look at the relationship of current and life-time earnings or Chetty *et al.*(2014) who look at the relationship of parental and child earnings show that measures using wages at the age of 30 are rather stable predictors of life-time earnings or intergenerational mobility, respectively. Finally, slightly less than 60 percent of the internal migrants change the industry in which they work when they move.

3. Historical Dialect Distance Between German Counties

Our proxy for cultural distance is based on historical dialect data of German localities. This unique source of data is taken from the language survey conducted for the Linguistic Atlas of the German Empire (*Sprachatlas des Deutschen Reichs*; data exploration between 1879 and 1888). Under the direction of the linguist Georg Wenker, pupils in more than 45,000 German schools were asked to translate 40 German sentences (more than 300 words) into their local dialect.[7] One of the prominent results of this project was the finding of a total of 66 prototypical characteristics of pronunciation and grammar that Wenker and his successors determined during a long lasting evaluation process (cf. Wrede *et al*. 1927). These characteristics are most relevant for structuring the German-language area.

Based on these prototypical characteristics, Falck *et al*. (2012) construct a dialect similarity matrix across all German counties (cf. Lameli 2013 for more details). For each county, the specific dialect is identified by the individual realization of the prototypical characteristics. The similarity between any two German counties is then quantified by counting the relative frequency of co-occurrences between any two profiles. We use this measure for further analysis, but, as we are dealing with the concept of cultural distance in this paper, we convert it from a similarity matrix into a distance matrix. The resulting dialect distance matrix across all counties has a dimension of 439×439, with elements ranging between 0 (linguistic identity) to 1 (maximum linguistic distance). To illustrate, Figure 1 shows the dialect distance of all other counties to the city of Worms (Rhineland-Palatinate). The figure reveals that dialect distance is low for counties to the east, west, and north of Worms, but high for counties in to the south of Worms.

<<Figure 1 about here>>

Panel B of Table 1 shows the result from merging the dialect distance matrix with the information on county of origin and destination of the internal migrants in the IAB Employment Panel. On average, an internal migrant moves 200 kilometers and experiences 0.372 in cultural distance by doing so.

[7] The results are available in the form of phonetic protocols for each school, cf. <http://www.regionalsprache.de>.

4. Estimation Strategy

To deal with unobserved self-selection into different locations, we adopt an estimation strategy from the literature on the effects of training programs on wages (e.g., Ashenfelter and Card 1985, LaLonde 1986). The basic idea in this strand of literature is to use pre-treatment wages to control for unobserved selection into programs. Comparing individuals with similar pre-treatment wage profiles should mitigate the selection problem. McKenzie *et al.* (2010) evaluate the transferability of this type of estimation strategy in the context of gains to migration and demonstrate that it performs relatively well. Specifically, we estimate the following wage regression:

$$\begin{aligned} \log wage_{idt+1}^{indexed} = \alpha + \beta Dialect\,Distance_{sd} + \sum_{j=1}^{4} \gamma_j \log wage_{ist-j}^{indexed} \\ + \lambda Geographical\,Distance_{sd} + \phi X_{it-1} + \mu_t + \varepsilon_{idt+1} \end{aligned} \tag{3}$$

The log of the indexed wage received by internal migrant *i* in destination *d* in the quarter after the move, i.e., at *t+1*, is regressed on the dialect distance between the origin county *s* and destination county *d*. The coefficient of interest is β, which is the wage premium in percent for overcoming one unit in dialect distance. The identification assumption under which the coefficient β would report the causal effect of dialect distance on wage after the move requires that the dialect distance is not correlated with unobserved individual characteristics. By including the last four quarterly wages before the move, $\sum_{j=1}^{4} \gamma_j \log wage_{ist-j}^{indexed}$, we control for unobserved self-selection into different locations. We further control for geographic distance between *s* and *d*.[8] We also control for gender, education (five dummies), experience (and its square), and a dummy indicating an industry change accompanying the move.[9] The quarter fixed effects μ_t capture all quarter-specific time shocks. Finally, ε_{idt+1} is an idiosyncratic error term.[10]

5. Wage Premium for Overcoming Dialect Distance

Table 2 shows our baseline results. The sample is restricted to the internal migrants' first move.[11] Column (1) shows the association between dialect distance and post-migration log indexed wage conditional on geographical distance and quarter-year fixed effects. Dialect distance is in fact positively correlated with post-migration log indexed wage. Interestingly,

[8] In a further specification, we also control for a higher polynomial function of geographic distance.
[9] Instead of using a dummy for the industry change, we can also use industry fixed effects. The results do not change.
[10] We use robust standard errors throughout the paper. In various robustness checks, we clustered standard errors at various levels. However, clustering at the origin county *x* destination county, the origin county, or the destination county yield very similar standard errors.
[11] We analyze multiple-time movers in Section 7.

conditional on dialect distance, the geographical distance enters negatively, which means that long-distance moves are associated with lower wages. In column (2), we add the last four quarterly pre-migration log indexed wages to control for unobserved self-selection into different locations. The coefficient on dialect distance drops by almost a factor of four and almost all pre-migration wages are highly significant predictors of post-migration wages. This indicates that self-selection is indeed a serious issue and neglecting pre-migration wages in the regression would lead to an upwardly biased estimate of the coefficient on dialect distance. After controlling for pre-migration wage profiles, adding further control variables in Column (3) hardly change the picture. The coefficient decreases slightly to 0.075 but is still highly significant. Thus, an increase in dialect distance by one standard deviation (about 0.2) increases the post-migration indexed wage by about 1.5 percent.

<< Table 2 about here >>

In Column (4) of Table 2, we provide an alternative specification where we explain the log wage without index and control for the log rental price in the destination and the source county on the right-hand side of the regression. The log rental price of the destination county is a significant predictor of post-migration log wages. Thus, wages in areas with high rental rates also relatively high. The rental price in the source county is not associated with the post-migration wage. The coefficient on the dialect distance decreases in this specification but is still in the same ballpark as the baseline specification. Increasing dialect distance by one standard deviation increases the post-migration wage by almost 1 percent.[12]

We can evaluate the magnitude of the wage premium needed to compensate for one standard deviation in dialect distance from different perspectives. The indexed wages of internal migrants increase on average by about 1.73 percent from the quarter before the move to the first quarter after the move. This implies that the wage premium necessary to compensate for one standard deviation in dialect distance is about 87 percent of the average wage gain from internal migration. The nominal wage increases from the quarter before the move to the quarter after the move by 4.32 percent. Thus, in nominal terms, the wage premium has to be 23 percent. The effect size of 1 percent per standard deviation is also sizeable when compared to the most recent (2013) collective wage agreements in Germany. In the public sector they agreed upon an increase of 2.65 percent in nominal wages (ver.di 2013) and in manufacturing they negotiated an increase of 3.4 percent in nominal wages (IG Metall 2013).

Figure 2 shows an add-variable binned scatterplot where we only use the variation in the post-migration indexed wage that is left over after taking account of the geographic distance, quarter-year fixed effects, and four quarterly pre-migration wages (blue) or conditional on the full control set (red). The conditional post-migration indexed wage is than

[12] The results are comparable when we use a price index instead of rental rates. However, we think that rental rates are able to capture amenities much better than price levels.

plotted against on the residual dialect distance after taking out all variation that is due to the full control set.[13] We can see that there is almost a linear relationship between the residual dialect distance and the conditional post-migration indexed wage once the dialect distance has crossed the 10th percentile (first two bins).

<< Figure 2 about here >>

6. Robustness Checks

We perform a couple of robustness checks. Table 3 restricts the sample to various subgroups. In Column (1), we look only at movers who are 30 or younger. There are two reasons for this robustness check. First, we want to check whether age plays a crucial role and second, younger people are more likely to resettle for the first time. Thus, we should be more likely to capture them when they leave their area where they have grown up for the first time. The coefficient is large and still highly significant, indicating that young movers act in a similar way then older movers.[14]

We are worried that the effect that we observe cluster only around large agglomerations. To check this, we exclude in Column (2) of Table 3 the five largest cities in Germany (Berlin, Hamburg, Munich, Cologne, and Frankfurt) as destination and source counties. Even though these cities account for almost a quarter of the mover sample, the coefficient of the regression stays virtually the same.[15]

Unfortunately, we do not know where the individuals in our sample were born and socialized, raising the concern that a migrant might not be attached to the county he or she left.[16] In this case, however, we would not expect to see any effect of cultural distance on post-migration wages. Thus, our baseline results should indicate a lower bound of the effect of cultural distance on migration wage gains.

To get some sense of the extent to which we underestimate the true effect of cultural distance on migration wage gains, we restrict the sample to those internal migrants who have not changed place of work or residence for a reasonable period before the move. Living in a region for a longer period can have the result that the person becomes more attached to that county than to the former home county (Burchardi and Hassan 2013). Given that our panel covers nine years, we decided to restrict our analysis to those 1,815 individuals who resided

[13] The figure shows a binned scatterplot where the residual dialect distance is binned into 20 equal-sized bins. The mean of each bin is then computed and plotted against the corresponding mean of the conditional post-migration indexed wage.

[14] Schwartz (1973) argues that the interaction of geographic distance with age should indicate the importance of the psychic cost of migration. Therefore, we also interacted geographic distance with age. In this specification (not shown), the interaction is not significant and the effect of dialect dissimilarity remains unchanged.

[15] In other specifications, we include dummies for moving from East to West Germany and for moving from West to East Germany or for changing the state. The results are not affected by these dummies.

[16] In particular, the migration flows in the aftermath of World War II (e.g., refugees, ethnic Germans, etc.) might have substantially reshuffled the German population with respect to cultural roots.

and worked in the origin county for at least seven years and then moved to a new destination during the last two years of our panel. The result of this procedure is shown in Column (3) of Table 3. The coefficient of dialect distance almost doubles. This is more evidence for our argument that the baseline effect is more of a lower bound and that being attached to a certain area for a longer period increases the cost of moving.

The last check in Column (5) introduces a couple of bilateral county controls which might affect migration decisions or could have led to a convergence of income levels. Thus, we include the log difference in slope, the historical rail distance, a dummy for a different religion, the difference in share Catholics, the difference in the historical industry structure, and the difference in the current industry structure. None of these controls change the coefficient on dialect distance significantly.

A crucial assumption is that the four quarterly pre-migration wages captures sufficiently the unobserved ability of the migrant. Table 4 introduces more demanding specification by conditioning on more pre-migration wages. Columns (1) to (8) include up to two more years of quarterly wages. However, the coefficient on dialect distance stays comparable in size. Column (9) changes the setup and includes the wages in yearly intervals. The coefficient on dialect distance increases which indicates that the quarterly wages capture the selection of internal migrants better than the yearly wages.[17]

<< Table 4 about here >>

The next concern is the possibility that dialect distance only captures non-linearities in geographic distance. For this reason, Table 5 shows several specifications by including non-linear geographic distance measures. Column (1) replicates the baseline regression. Column (2) includes the geographic distance of power two and three. The coefficient on dialect distance increases which indicates that it does not capture strong non-linearities in geographical distance.

<< Table 5 about here >>

However, it could be that the geographical distance is an insufficient distance proxy for the dialect distance. Therefore, we use in Column (3) and (4) of Table 5 the travel distance by car in minutes between counties as an alternative geographical index. However, the travel time could be affected by dialect distance as well because it is very likely that transportation hubs and networks have developed along these lines. Column (3) shows the specification where we include the travel distance instead of the geographic distance. The coefficient decreases but stays highly significant. When we include the travel distance to the power two and three, we see that the coefficient increases slightly again. However, the results of this exercise indicate

[17] Using yearly averages instead of yearly wages leads to very similar results.

that there are, if at all, only minor nonlinear effects of geographic distance that is picked up by dialect distance.

Table 6 uses alternative measures of dialect distance to check the robustness of our results. As the most important linguistic difference of German dialects is the distinction of Low German dialects (northern part of Germany) and High German dialects (southern part), we construct a dummy that substantiate the particular localization of the counties and test for the movements within the two larger areas of Low German and High German. Column (1) of Table 6 include a dummy for moving from a county which High German to a county with Low German, a dummy for moving from a county with Low German to a county with High German, and a dummy for moving from a county with High German to a county with High German. The omitted category is moving from a county with Low German to a county with Low German. The results show that the effect of dialect distance remains robust when testing for the north-south distinction. We find, however, a slight north-south divide indicating the relevance of a categorical conceptualization of cultural space. While the wages are positively affected by migration from south to north, we find a negative association for migration from north to south.

<< Table 6 about here >>

In Columns (2) and (3), we test another measure of dialect distance which is based on an Euclidean dialect distance between counties. This measure shows highly similar results to our baseline dialect distance measure. Thus, increasing the Euclidean dialect distance by one standard deviation (2.214), increases the post-migration indexed wage by 1.79 percent (Column (2)) or 1.48 percent (Column (3)).

So far, we use a metric measure of dialect distance. It might be, however, that cultural space is not only dependent on gradual differences but categorical ones. That is, that the decision to move could be due to a difference between, e.g., “Swabian” and “Bavarian” as such and not to the actual gradual difference between the counties within the Swabian and Bavarian region. To test for the impact of categorical differences between smaller regions on migration we use a classification introduced by Lameli (2013) that captures the most prominent thirteen dialect areas within Germany. It results from bootstrapped hierarchical cluster analysis, based on the measurement of linguistic similarity of German counties. Column (4) of Table 6 provides the results. The coefficient is positive and highly significant. A one standard deviation in the dialect distance by language area (1.0433) leads to 1.11 percent higher post-migration indexed wages.

7. Effect Heterogeneity

The question arises as to whether there is a group of individuals who are driving the baseline results. To answer this question, we first spilt the sample by (i) gender, (ii) education, (iii) education x gender, and (iv) distance of the move. The results are summarized in Table 7.

<< Table 7 about here >>

Table 7, Panel A shows that the wages of men are more responsive to culture than those of women. Possibly this is because in most families the male adult is the household head and his place of work determines to a large extent where the family lives. We also see that low- and medium-qualified migrants find culture more of a barrier to migration than do higher qualified migrants.[18] However, the difference is not very large between the two groups. Panel B shows that within the group of men, it is again the group of lower qualified migrants that show a larger coefficient. But the differences to the lower skilled men are not significant. The results for women are insignificant again and the coefficient for lower qualified women is slightly higher than for higher qualified. Panel C reveals that the effect is mainly coming from shorter-distance moves, i.e., moves less than 300 km from the former home county. Thus, the wage increases from moving to a more cultural distant county are not driven by long-distance moves as one might have first expected. We also looked at the subsample of internal migrants who also switch occupation along with the move in Panel D. Compared to occupational stayers, switchers are compensated more for their move to a dialect-dissimilar county.

We also analyze in more detail the 567 two-time movers in our sample.[19] Recall that the total time period under analysis is nine years, meaning that every second move occurs within a relatively short time window. For the second move, we use the dialect distance and geographic distance between the origin county of the first move and the destination county of the second move. This should mimic the potential direct move to the destination county in the second move. All other control variables (quarterly pre-treatment wages, education, experience, age, etc.) are taken from the second move.

Interestingly, almost 34 percent (194 migrants) of the two-time movers in our sample return in the second move to exactly the same county from which they came. However, only 32 of the 194 repatriates return to the same firm.[20]

[18] The group of low- and medium-qualified people consists of people with a degree from the lowest and middle academic track with and without vocational education and training (VET). We also include people for whom level of education is unknown. However, the picture does not change by omitting this group. The group of high-qualified people consists of people with a degree from the highest academic track or with an university degree.

[19] There are some individuals who moved more than two times, but this group is too small for an in-depth investigation.

[20] "Repatriates" are those who move to their previous county of residence and again work in their previous county of work. Thus, migrants who return to their previous home county but work in a different county than before are not considered as repatriates.

Table 8 shows the results of the two-time mover analysis by first and second move and by timing of the move, i.e. whether the migrant has moved another time after or before eight quarters (two years). Panel A shows the results with repatriates and Panel B shows the results by excluding them. The first move shows, independent of the timing, that the coefficients are larger than in the baseline sample.[21] This indicates that this particular group of people, we know that they will move within the next nine years for another time, value culture highly. The second move is more interestingly. We see that the coefficient for those who moved another time within eight quarters is almost seven times as large as the baseline coefficient. The above findings lead us to view these two-time movers as members of a selected group who made a wrong choice about where to live and work in the first move and are now willing to sacrifice a lot more money in return for a more familiar environment.

<< Table 8 about here >>

8. Persistence of Wage Premium

We now turn to the question of whether the initial effect directly after the move is persistent over time. To this end, we look at wage growth after the first move estimating the regression in Equation (2).

$$\log wage^{real}_{idt+k} - \log wage^{real}_{idt+1} = \alpha + \beta Dialect\,Distance_{sd} + \gamma \log wage^{real}_{ist+1} + \lambda Geographical\,Distance_{sd} + \phi X_{it-1} + \mu_t + \varepsilon_{idt+1} \quad (2)$$

Conditional on the logged initial wage level after the move, we regress wage growth from period *t+1*, i.e., the first quarter after the move, to period *t+k*, on dialect distance. Thereby, *k* takes a maximum value of 32 (quarters), i.e., we analyze wage growth within a maximum of eight years after the move. Note that by extending the growth period of analysis year by year, the number of internal migrants remaining in the sample drops significantly, until finally there are less than 700 internal migrants included in the analysis of eight-year post-move wage growth. All other control variables remain equivalent to the baseline model. Due to a "catching-up" process, we expect that migrants who moved to culturally more dissimilar counties will show lower wage growth rates. Table 9 shows the results for the three- to six-year wage growth rates. The coefficient on the logged initial wage level after the move shows that internal migrants with initially higher wages after the move in general experience lower wage growth in the future. However, dialect distance is not significantly associated with future wage growth. Thus, we conclude that the initial wage sacrifice is persistent over time.

<< Table 9 about here >>

[21] However, the large standard error does not allow us to rule out positive coefficients as well.

9. Conclusion

In this paper, we monetize the psychic cost of migration by combining administrative social security panel data with a proxy for cultural difference that is constructed from unique data on historical dialect distance between German counties. Internal migrants demand a wage premium of about 2 percent for a one standard deviation in dialect distance. Compared to the general wage gain associated with internal migration and compared to general wage increases negotiated in the recent collective agreement, this wage premium, which is arguably a lower-bound estimate, is economically substantive and persistent over time. Our results imply that analyses of rate of return to migration that do not consider the psychic cost of migration overestimate the rate of return to resources allocated to migration.

Literature

Ashenfelter, Orley C. and David Card (1985), Using Longitudinal Structure of Earnings to Estimate the Effect of Training Programs. *Review of Economics and Statistics*, 67(4), 648–660.

Barrett, Alan and Irene Mosca (2013), The Psychic Costs of Migration: Evidence from Irish Return Migrants, *Journal of Population Economics*, 26(2), 483–506.

Burchardi, Konrad B. and Tarek A. Hassan (2013), The Economic Impact of Social Ties: Evidence from German Reunification. *Quarterly Journal of Economics*, 128(3), 1219–1271.

Card, David, Jörg Heining, and Patrick Kline (2013), Workplace Heterogeneity and the Rise of West German Wage Inequality. *Quarterly Journal of Economics*, 128(3), 967–1015.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner (2014). Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States. NBER Working Paper 19844.

Dahl, Michael S. and Olav Sorenson (2010), The Migration of Technical Workers. *Journal of Urban Economics,* 67(1), 33–45.

Dustmann, Christian, Johann Ludsteck, and Uta Schönberg (2009). Revisiting the German Wage Structure. *Quarterly Journal of Economics*, 124(2), 843–881.

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

Federal Statistical Office (2013), Bevölkerung und Erwerbstätigkeit, Fachserie 1, Reihen 1.2 und 1.3.

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

Guiso, Luigi, Paula Sapienza, and Luigi Zingales (2006), Does Culture Affect Economic Outcomes? *Journal of Economic Perspectives*, 20(2), 23–48.

Haider, Steven and Gary Solon (2006), Life-Cycle Variation in the Association between Current and Lifetime Earnings. *American Economic Review*, 96(4), 1308–1320.

IG Metall (2013), Collective Wage Agreement in the Manufacturing Sector 2013, http://www.igmetall.de/SID-49F5A1E2-969B8B98/metall-tarifrunde-2013-bayerisches-ergebnis-in-allen-11891.htm

LaLonde, Robert J. (1986), Evaluating the Econometric Evaluations of Training Programs with Experimental Data, *American Economic Review,* 76(4), 604–620.

Lameli, Alfred (2013), *Strukturen im Sprachraum. Analysen zur arealtypologischen Komplexität der Dialekte in Deutschland*. De Gruyter, Berlin, Boston.

McKenzie, David, John Gibson, and Steven Stillman (2010), How Important is Selection? Experimental vs. Non-Experimental Measures of the Income Gains from Migration, *Journal of the European Economic Association*, 8(4), 913–945.

Schmucker, Alexandra and Stefan Seth (2009), BA-Beschäftigtenpanel 1998–2007 Codebuch, FDZ Datenreport 01/2009, Bundesagentur für Arbeit Nürnberg.

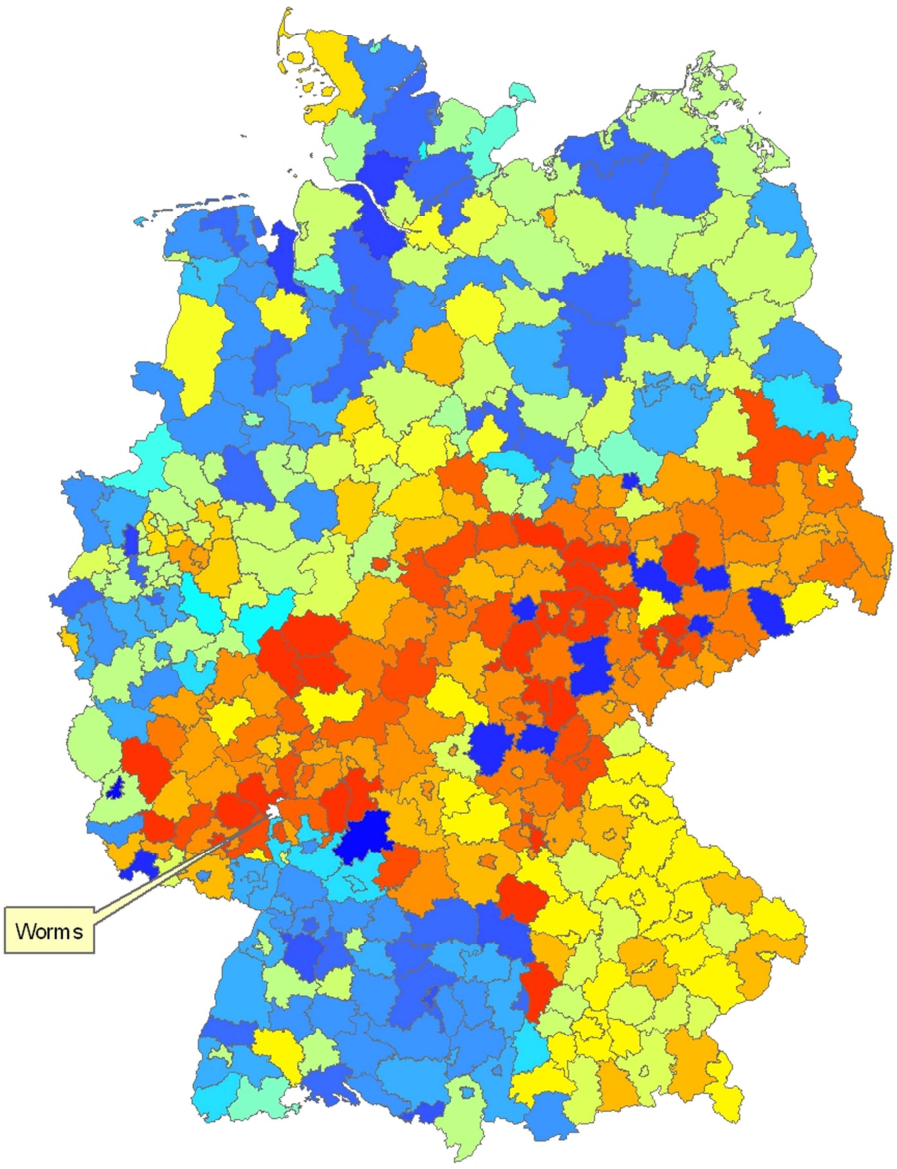
Schwartz, Aba (1973), Interpreting the Effect of Distance on Migration, *Journal of Political Economy*, 81(5), 1153–1169.

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

ver.di (2013), Collective Wage Agreement in the Public Sector 2013, http://www.verdi.de/themen/geld-tarif/tarifrunde-oeffentlicher-dienst-der-laender-2013/++co++a5c366fa-784c-11e2-8d18-52540059119e

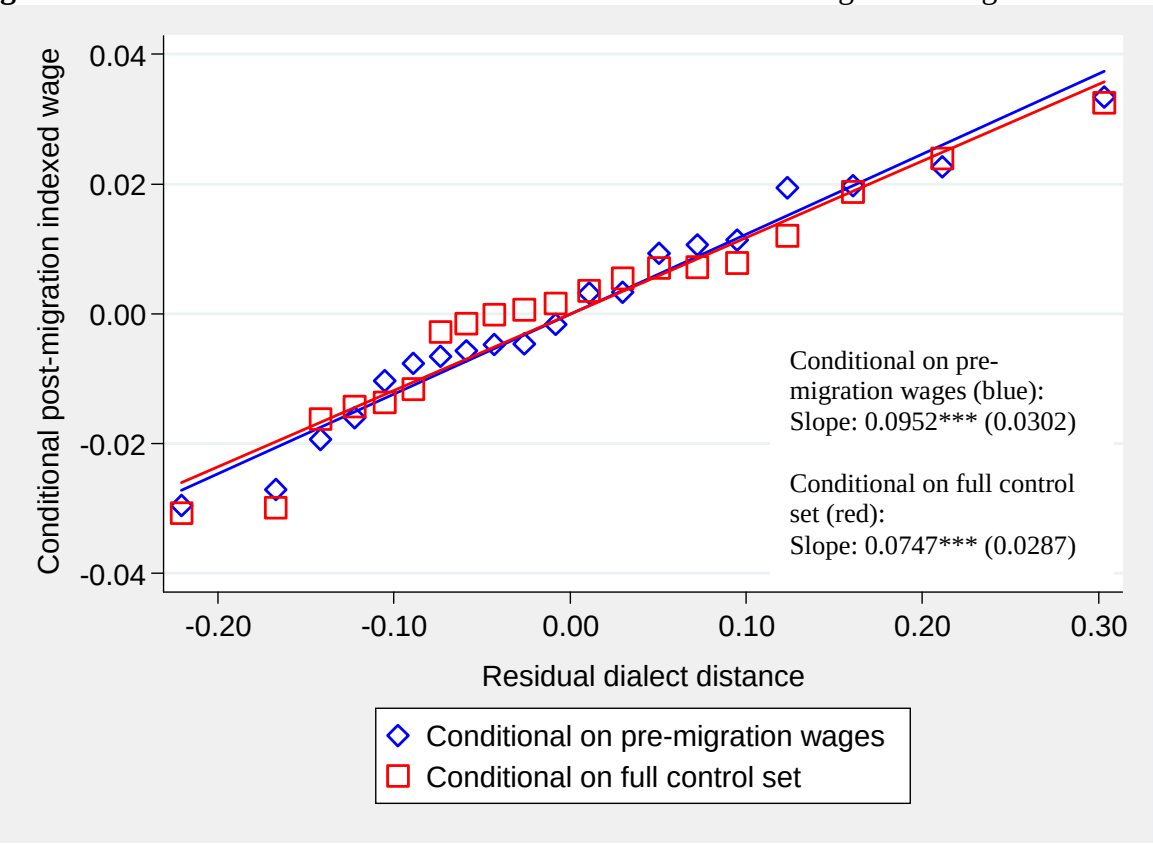
Wrede, Ferdinand, Walther Mitzka, and Bernhard Martin (1927), Deutscher Sprachatlas. Auf Grund des von Georg Wenker begründeten Sprachatlas des Deutschen Reichs. Elwert, Marburg.

Figure 1: Dialect Distance —The Case of Worms



Notes: The figure shows dialect distance of all districts to the reference point Worms. Degrees of dialect distance (from highest to lowest) are indicated by: blue, green, yellow, red, dark to light gray.

Figure 2: Added-Variable Plot of Dialect Distance and the Post-Migration Wage



Notes: The figure shows a binned scatterplot of conditional post-migration indexed wages on residual dialect distance. The conditional post-migration indexed wages are obtained from residuals from regressions on the geographical distance, quarter-year fixed effects and the last four quarterly pre-migration wages (blue) or additionally on education dummies (4), male, experience, experience squared, and an industry change dummy (red). The residual dialect distance is obtained from residuals from regressions on the geographical distance, quarter-year fixed effects, the last four quarterly pre-migration wages, education dummies (4), male, experience, experience squared, and an industry change dummy. The figure is constructed by binning dialect distance into five-percentile point bins (so that there are 20 equal sized bins). Coefficients and robust standard deviation in parentheses are obtained from the regressions on the micro data. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 1: Descriptive Statistics

	Total sample		Dialect distance			
			below median		above median	
Variable	**Mean (SD)**	**Min Max**	**Mean (SD)**	**Min Max**	**Mean (SD)**	**Min Max**
Panel A: Wage data						
Wage (t+1)	2,659	213	2,547	213	2,766	213
	(1,157)	5,250	(1,134)	5,250	(1,168)	5,250
Wage (t-1)	2,549	202	2,449	202	2,644	222
	(1,148)	5,250	(1,122)	5,250	(1,165)	5,250
Wage (t-2)	2,526	201	2,428	202	2,619	201
	(1,159)	5,250	(1,130)	5,221	(1,179)	5,250
Wage (t-3)	2,492	203	2,397	203	2,583	204
	(1,170)	5,250	(1,141)	5,250	(1,190)	5,250
Wage (t-4)	2,466	201	2,375	203	2,553	201
	(1,174)	5,250	(1,145)	5,178	(1,195)	5,250
Indexed wage (t+1)	4,703	280	4,617	280	4,784	283
	(2,150)	14,173	(2,152)	13,658	(2,146)	14,173
Indexed wage (t-1)	4,623	242	4,472	242	4,767	414
	(2,119)	13,448	(2,101)	13,448	(2,127)	12,468
Indexed wage (t-2)	4,581	242	4,432	242	4,724	306
	(2,135)	13,448	(2,108)	13,448	(2,150)	12,468
Indexed wage (t-3)	4,525	203	4,380	203	4,664	375
	(2,157)	13,448	(2,133)	13,448	(2,171)	12,468
Indexed wage (t-4)	4,482	203	4,347	203	4,611	325
	(2,159)	13,448	(2,140)	13,448	(2,169)	13,008
Panel B: Distance data						
Dialect distance	0.372	0	0.189	0	0.548	0.379
	(0.207)	0.833	(0.103)	0.364	(0.104)	0.833
Geographical distance (km)	200	1	76	1	318	15
	(170)	818	(72)	595	(150)	818
Panel C: Individual characteristics						
Lowest and middle academic track, without VET	0.019	0 1	0.021	0 1	0.017	0 1
Lowest and middle academic track, with VET	0.523	0 1	0.572	0 1	0.477	0 1
Highest academic track	0.144	0 1	0.135	0 1	0.152	0 1
University	0.300	0 1	0.260	0 1	0.339	0 1
Education unknown	0.014	0 1	0.012	0 1	0.015	0 1
Male	0.557	0 1	0.570	0 1	0.545	0 1
Age	32.038	18	31.965	18	32.108	18
	(8.049)	63	(8.307)	62	(7.794)	63
Experience	11.687	0	11.790	0	11.587	0
	(8.083)	43	(8.335)	43	(7.835)	43
Industry change	0.588	0 1	0.578	0 1	0.598	0 1
Observations	9,090		4,444		4,646	

Notes: Summary statistics are based on the baseline sample. Only the first observed move is considered for individuals who moved several times. We have full information on 567 individuals, or 6.2 percent, who moved two times during our time period. Wage data and data on individual characteristics are drawn from the IAB Employment Panel. The distance data are from Falck *et al*. (2012). Standard deviations are not computed for dummy variables. The variable *t* indicates the timing of the move. *t+1* denotes the first observation after the move, *t=0* denotes the move, and *t-1* denotes the quarter before the move. *Experience* represents potential labor market experience and is computed by *Age – 6 – years of schooling. Years of schooling* is assumed to be equal to 10 years for *lowest and middle academic track without VET*, 13 years for *lowest and middle academic track with VET*, 13 years for *highest academic track without VET*, 15 years for *highest academic track with VET*, 17 years for *university*, and 10 years for *education unknown.* We merged *highest academic track without VET* and *highest academic track with VET* into one education category.

Table 2: Dialect Distance and Post-Migration Wages

Dependent variable:	*Log indexed wage (t+1)*			*Log wage (t+1)*
	(1)	(2)	(3)	(4)
Dialect distance	0.3872***	0.0952***	0.0747***	0.0486**
	(0.0389)	(0.0302)	(0.0287)	(0.0234)
Geographical distance (km)	-0.00035***	-0.00012***	-0.00013***	-0.000002
	(0.00005)	(0.00004)	(0.00003)	(0.000029)
Log indexed wage (t-1)		0.342***	0.2608***	
		(0.0272)	(0.0256)	
Log indexed wage (t-2)		0.1137***	0.088***	
		(0.0334)	(0.0301)	
Log indexed wage (t-3)		-0.0134	-0.0098	
		(0.0376)	(0.0334)	
Log indexed wage (t-4)		0.2000***	0.1509***	
		(0.0293)	(0.0267)	
Log wage (t-1)				0.3799***
				(0.0273)
Log wage (t-2)				0.0984***
				(0.0307)
Log wage (t-3)				0.0178
				(0.035)
Log wage (t-4)				0.0897***
				(0.0286)
Log rental price (t+1)				0.1984***
				(0.0144)
Log rental price (t-1)				-0.0216
				(0.0158)
Lowest and middle academic track, with VET			0.0869***	0.0635**
			(0.0332)	(0.029)
Highest academic track			0.2002***	0.1685***
			(0.0346)	(0.0304)
University			0.3537***	0.2939***
			(0.0344)	(0.0305)
Education unknown			0.167***	0.1114***
			(0.0473)	(0.0408)
Male			0.0928***	0.0765***
			(0.0081)	(0.007)
Experience			0.0124***	0.0059***
			(0.0017)	(0.0015)
Experience squared x 10^{-4}			-2.7144***	-1.2637***
			(0.5098)	(0.4391)
Industry change			-0.0542***	-0.0366***
			(0.0073)	(0.0061)
Quarter-year fixed effects	YES	YES	YES	YES
Observations	9,090	9,090	9,090	9,090
R^2	0.0225	0.4399	0.4994	0.4997

Notes: The log indexed wage is the gross wage divided by the index of the rental rate. The log rental price at t+1 is the average rental price in the destination county and the log rental price is the average rental price in the source county. Only the first observed move is considered for individuals who moved several times. The omitted education category is *Lowest and Middle Academic Track, without VET*. Robust standard errors in parentheses. *t = 0* denotes the time of the move. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: Robustness Checks

	Dependent variable: Log indexed wage (t+1)			
	(1)	(2)	(3)	(5)
	Young movers	5 largest cities	7 years at origin	Bilateral controls
Dialect distance	0.093**	0.0741**	0.1295**	0.0933***
	(0.0404)	(0.0303)	(0.0653)	(0.0292)
Geographical Distance (km)	-0.000158***	-0.000149***	-0.000158**	-0.00015***
	(0.000049)	(0.000038)	(0.000078)	(0.000037)
Log indexed wage (t-1)	0.2245***	0.286***	0.3601***	0.2762***
	(0.0285)	(0.0297)	(0.0623)	(0.0257)
Log indexed wage (t-2)	0.0965***	0.0903**	-0.0103	0.0807***
	(0.0311)	(0.0357)	(0.0783)	(0.0302)
Log indexed wage (t-3)	-0.0463	0.0304	0.0937	-0.0129
	(0.0316)	(0.0403)	(0.0586)	(0.0335)
Log indexed wage (t-4)	0.1396***	0.1224***	0.0692*	0.1538***
	(0.027)	(0.0317)	(0.0363)	(0.0265)
Lowest and middle academic track, with VET	0.0618	0.0422	0.053	0.0871***
	(0.0429)	(0.0321)	(0.0748)	(0.0327)
Highest academic track	0.1795***	0.1383***	0.1549**	0.1992***
	(0.0444)	(0.0341)	(0.0773)	(0.0341)
University	0.3592***	0.3092***	0.3271***	0.3504***
	(0.0453)	(0.0339)	(0.0768)	(0.0339)
Education Unknown	0.094	0.1164**	0.162*	0.1691***
	(0.0678)	(0.0456)	(0.0932)	(0.0473)
Male	0.0784***	0.0919***	0.1335***	0.0946***
	(0.011)	(0.0089)	(0.0189)	(0.008)
Experience	0.0097	0.0075***	0.0086**	0.011***
	(0.0069)	(0.0018)	(0.0038)	(0.0017)
Experience squared x 10^{-4}	3.6725	-1.5633***	-2.2194**	-2.4079***
	(5.8789)	(0.5601)	(1.1222)	(0.5058)
Industry change	-0.0432***	-0.0484***	-0.0527***	-0.0531***
	(0.0104)	(0.0079)	(0.017)	(0.0072)
Quarter-year fixed effects	YES	YES	YES	YES
Observations	4,384	6,946	1,815	9,090
R^2	0.3948	0.5395	0.5327	0.5107

Notes: Column (1) keeps only movers below the age of 30. Column (2) drops the five largest cities (Berlin, Hamburg, Munich, Cologne, Frankfurt) as destination and source counties. Column (3) conditions the sample on having lived at least seven years in the county. Column (4) keeps all movers who change the firm most likely for the first time (at least in the sample period, we observe). Column (5) includes several bilateral controls: log difference in slope, historical rail distance, different religion dummy, difference in share Catholics, difference in historical industry structure, and the difference in the current industry structure. Only the first observed move is considered for individuals who moved several times. $t = 0$ denotes the time of the move. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: Adding more/other Pre-Migration Wages

	Dependent variable: Log indexed wage (t+1)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dialect distance	0.0790***	0.0637**	0.0629**	0.0550*	0.0666**	0.0821**	0.0916**	0.0846**	0. 1106***
	(0.0294)	(0.0299)	(0.0308)	(0.0319)	(0.0334)	(0.0347)	(0.0368)	(0.0386)	(0.0379)
Log indexed wage (t-4)	0.0765**	0.0422	0.0298	-0.0005	0.0084	0.015	0.0178	0.0136	0.2709***
	(0.0343)	(0.0386)	(0.0407)	(0.0441)	(0.0447)	(0.0474)	(0.0498)	(0.0543)	(0.0211)
Log indexed wage (t-5)	0.0777***	0.0706*	0.0731	0.0668	0.0563	0.04	0.0316	0.0331	
	(0.0229)	(0.0426)	(0.0445)	(0.0498)	(0.0517)	(0.0532)	(0.0571)	(0.0604)	
Log indexed wage (t-6)		0.0211	-0.0415	-0.042	-0.0507	-0.0334	-0.0384	-0.0466	
		(0.0318)	(0.0472)	(0.053)	(0.0572)	(0.0585)	(0.0632)	(0.0663)	
Log indexed wage (t-7)			0.0746**	0.0602	0.0583	0.067	0.0682	0.0859	
			(0.0288)	(0.0512)	(0.0556)	(0.0597)	(0.0637)	(0.0665)	
Log indexed wage (t-8)				0.0291	-0.0369	-0.0723	-0.0652	-0.0914	0.0559**
				(0.0321)	(0.0506)	(0.0577)	(0.0635)	(0.0694)	(0.0231)
Log indexed wage (t-9)					0.0932***	0.0643	0.0616	0.0612	
					(0.0307)	(0.0422)	(0.0462)	(0.0502)	
Log indexed wage (t-10)						0.0577*	0.0454	0.0659	
						(0.0298)	(0.04)	(0.0443)	
Log indexed wage (t-11)							0.0225	-0.0123	
							(0.0254)	(0.0366)	
Log indexed wage (t-12)								0.0317	0.0851***
								(0.0311)	(0.0177)
Quarter-year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
Quarterly pre-migration wages (t-1,…,t-3)	YES	YES	YES	YES	YES	YES	YES	YES	NO
Observations	8,517	7,949	7,392	6,845	6,225	5,692	5,099	4,681	5,411
R^2	0.5048	0.5137	0.522	0.5225	0.5208	0.5283	0.5276	0.5293	0.4955

Notes: Only the first observed move is considered for individuals who moved several times. *Controls:* geographical distance, education, male, age, age squared, experience, experience squared, and industry change. Robust standard errors in parentheses. $t = 0$ denotes the time of the move. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: Non-Linearities in Distance

	Dependent variable: Log indexed wage (t+1)			
	(1)	(2)	(3)	(4)
Dialect distance	0.0747***	0.0846***	0.0603**	0.0640**
	-0.0287	-0.0324	(0.0291)	(0.0315)
Geographical distance (km)	-0.00013***	0.00002		
	-0.00003	-0.00018		
Travel distance			-0.00014***	0.00029
			(0.00005)	(0.00025)
Geographical distance (km) squared x 10^{-5}		-0.0924		
		-0.064		
Geographical distance (km) cubic x 10^{-5}		0.00012*		
		-0.00007		
Travel distance (min) squared x 10^{-3}				-0.002**
				(0.001)
Travel distance (min) cubic x 10^{-5}				0.0003**
				(0.0001)
Log indexed wage (t-1)	0.2608***	0.2608***	0.2614***	0.2614***
	-0.0256	-0.0255	(0.0256)	(0.0255)
Log indexed wage (t-2)	0.088***	0.0874***	0.0882***	0.0868***
	-0.0301	-0.03	(0.0301)	(0.0301)
Log indexed wage (t-3)	-0.0098	-0.0095	-0.0099	-0.0091
	-0.0334	-0.0334	(0.0335)	(0.0334)
Log indexed wage (t-4)	0.1509***	0.1514***	0.1512***	0.1514***
	-0.0267	-0.0267	(0.0268)	(0.0267)
Lowest and middle academic	0.0869***	0.088***	0.0865***	0.0868***
track, with VET	-0.0332	-0.0333	(0.0332)	(0.0333)
Highest academic track	0.2002***	0.2015***	0.1992***	0.2***
	-0.0346	-0.0347	(0.0346)	(0.0346)
University	0.3537***	0.3552***	0.3525***	0.3531***
	-0.0344	-0.0345	(0.0344)	(0.0344)
Education unknown	0.167***	0.1686***	0.1661***	0.1673***
	-0.0473	-0.0474	(0.0474)	(0.0474)
Male	0.0928***	0.0932***	0.093***	0.0936***
	-0.0081	-0.0081	(0.0081)	(0.0081)
Experience	0.0124***	0.0124***	0.0123***	0.0124***
	-0.0017	-0.0017	(0.0017)	(0.0017)
Experience squared x 10^{-4}	-2.7144***	-2.7138***	-2.7126***	-2.7249***
	-0.5098	-0.5089	(0.5097)	(0.509)
Industry change	-0.0542***	-0.0538***	-0.0542***	-0.0537***
	-0.0073	-0.0073	(0.0073)	(0.0073)
Quarter-year fixed effects	YES	YES	YES	YES
Observations	9,090	9,090	9,090	9,090
R^2	0.4994	0.4997	0.4991	0.4994

Notes: Only the first observed move is considered for individuals who moved several times. The omitted education category is *Lowest and Middle Academic Track, without VET*. Robust standard errors in parentheses. *t = 0* denotes the time of the move. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6: Alternative Dialect Measures

	Dependent variable: Logged indexed wage (t+1)			
	(1)	(2)	(3)	(4)
Dialect distance	0.0582*			
	(0.0307)			
Euclidean dialect distance		0.0081***	0.0067**	
		(0.0026)	(0.0027)	
Dialect distance by language area				0.0106**
				(0.005)
Moving from HG to LG	0.0847***		0.0849***	
	(0.0148)		(0.0143)	
Moving from LG to HG	-0.0955***		-0.095***	
	(0.0144)		(0.0141)	
Moving from HG to HG	-0.0425***		-0.0423***	
	(0.0095)		(0.0094)	
Geographical Distance (km)	-0.000131***	-0.000135***	-0.000142***	-0.000104***
	(0.000033)	(0.000032)	(0.000032)	(0.000031)
Log indexed wage (t-1)	0.2687***	0.2600***	0.268***	0.2619***
	(0.0252)	(0.0255)	(0.0252)	(0.0255)
Log indexed wage (t-2)	0.0866***	0.0882***	0.0868***	0.0874***
	(0.0298)	(0.0301)	(0.0297)	(0.0301)
Log indexed wage (t-3)	-0.0077	-0.01	-0.0079	-0.0095
	(0.0331)	(0.0334)	(0.0331)	(0.0335)
Log indexed wage (t-4)	0.1476***	0.151***	0.1476***	0.151***
	(0.0265)	(0.0267)	(0.0265)	(0.0267)
Lowest and middle academic track, with VET	0.0795**	0.0872***	0.0799**	0.0857***
	(0.0332)	(0.0332)	(0.0332)	(0.0332)
Highest academic track	0.1924***	0.2003***	0.1926***	0.1998***
	(0.0346)	(0.0346)	(0.0346)	(0.0345)
University	0.3445***	0.3537***	0.3446***	0.3532***
	(0.0344)	(0.0344)	(0.0344)	(0.0343)
Education Unknown	0.1652***	0.1671***	0.1653***	0.1664***
	(0.047)	(0.0473)	(0.0469)	(0.0474)
Male	0.0912***	0.0932***	0.0915***	0.0924***
	(0.008)	(0.0081)	(0.008)	(0.0081)
Experience	0.0126***	0.0124***	0.0126***	0.0125***
	(0.0017)	(0.0017)	(0.0017)	(0.0017)
Experience squared x 10^{-4}	-2.7804***	-2.7145***	-2.7791***	-2.7449***
	(0.5094)	(0.5094)	(0.509)	(0.51)
Industry change	-0.0516***	-0.0542***	-0.0516***	-0.0541***
	(0.0072)	(0.0073)	(0.0072)	(0.0073)
Quarter-year fixed effects	YES	YES	YES	YES
Observations	9,090	9,090	9,090	9,090
R^2	0.5093	0.4996	0.5095	0.4992

Notes: Moving from HG to LG indicates a move from a county in which mostly High German is spoken to a county in which mostly Low German is spoken. *Moving from LG to HG* indicates a move from a county in which mostly Low German is spoken to a county in which mostly High German is spoken. *Moving from HG to HG* indicates a move from a county in which mostly High German is spoken to a county in which mostly High German is spoken. The omitted category is *Moving from LG to LG*.

Table 7: Effect Heterogeneities

	Dependent variable: Logged indexed wage (t+1)			
	(1)	(2)	(3)	(4)
Panel A: Gender and education				
	Gender		Education	
	Men	Women	Low, medium	High
Dialect distance	0.1046***	0.0367	0.0844**	0.0669
	(0.0367)	(0.0451)	(0.0406)	(0.0410)
Observations	5,063	4,027	5,051	4,039
R^2	0.5057	0.4027	0.3615	0.4351
Panel B: Gender x education				
	Men		Women	
	Low, medium	High	Low, medium	High
Dialect distance	0.1143**	0.0934*	0.0519	0.0215
	(0.0548)	(0.0492)	(0.0598)	(0.0700)
Observations	2,727	2,336	2,324	1,703
R^2	0.3522	0.3872	0.3377	0.3457
Panel C: Geographical distance				
	< 200 km	< 300 km	≥ 200 km	≥ 300 km
Dialect distance	0.0984**	0.1179***	0.0534	-0.0279
	(0.0402)	(0.0356)	(0.0545)	(0.0702)
Observations	5325	6,575	3765	2,515
R^2	0.5155	0.5112	0.4862	0.4822
Panel D:Occupational change				
	Occupational information available	Occupational switchers	Occupational stayers	
Dialect distance	0.0806**	0.1154**	0.0423	
	(0.0313)	(0.0506)	(0.0383)	
Observations	7,337	3,479	3,858	
R^2	0.5089	0.4638	0.5757	

Notes: Low and medium education corresponds to the lowest and middle academic track with and without VET, plus unknown education. *High education* corresponds to the highest academic track plus university. Only the first observed move is considered for individuals who moved several times. All regressions include quarter-year fixed effects, four quarterly pre-migration wages, individual controls, and geographical distance controls. *Individual controls:* education, male, age, age squared, experience, experience squared, industry change. *Geographical distance controls:* geographical distance. Robust standard errors in parentheses. $t = 0$ denotes the time of the move. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 8: Two-Time Mover Analysis

Dependent variable: Logged indexed wage (t+1)				
	(1)	(2)	(3)	(4)
	First move		Second move	
	≥ 8 quarters	< 8 quarters	≥ 8 quarters	< 8 quarters
Panel A: With repatriates				
Dialect distance	0.2907*	0.1208	0.0363	0.5027***
	(0.1752)	(0.1881)	(0.1411)	(0.1559)
Observations	245	322	245	322
R^2	0.5813	0.5157	0.7002	0.5616
Panel B: Without repatriates				
Dialect distance	0.3145	0.2570	0.1621	0.6122***
	(0.2001)	(0.2693)	(0.1570)	(0.1932)
Observations	192	181	192	181
R^2	0.6224	0.5166	0.7074	0.6019

Notes: Repatriates are those individuals who move back to their county of origin in the second move. The *second move* uses the dialect and geographical distance between the source county of the first move and the destination county of the second move. All regressions include quarter-year fixed effects, four quarterly pre-migration wages, individual controls, and geographical distance controls. *Individual controls:* education, male, age, age squared, experience, experience squared, industry change. *Geographical distance controls:* geographical distance. $t = 0$ denotes the time of the move. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 9: Long-Run Effects

	Dependent variable: log indexed wage (t+k) - log indexed wage (t+1)							
	k = 4	*k = 8*	*k = 12*	*k = 16*	*k = 20*	*k = 24*	*k = 28*	*k = 32*
Dialect distance	0.0352*	0.0277	0.0293	0.0084	0.0066	0.0534	-0.0056	-0.0064
	(0.0190)	(0.0275)	(0.0365)	(0.0416)	(0.0478)	(0.0587)	(0.0809)	(0.1408)
Log indexed wage (t+1)	-0.1732***	-0.2713***	-0.3384***	-0.3566***	-0.3883***	-0.4002***	-0.4588***	-0.4801***
	(0.0139)	(0.0179)	(0.0206)	(0.0235)	(0.0269)	(0.0303)	(0.0392)	(0.0671)
Observations	8,209	6,872	5,875	4,910	4,076	3,142	1,923	698
R^2	0.0928	0.1226	0.1203	0.1317	0.1336	0.1274	0.1450	0.1628

Notes: Only the first observed move is considered for individuals who moved several times. All regressions include quarter-year fixed effects, individual controls, and geographical distance controls. *Individual controls:* education, male, age, age squared, experience, experience squared, industry change. *Geographical distance controls:* geographical distance. Robust standard errors in parentheses. *t = 0* denotes the time of the move. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.