DISCUSSION PAPER SERIES

IZA

IZA DP No. 8922

Cultural Biases in Migration: Estimating Non-Monetary Migration Costs

Oliver Falck
Alfred Lameli
Jens Ruhose

March 2015

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Cultural Biases in Migration: Estimating Non-Monetary Migration Costs

Oliver Falck
Ifo Institute, University of Munich (LMU) and CESifo

Alfred Lameli
Research Centre Deutscher Sprachatlas

Jens Ruhose
Ifo Institute and IZA

Discussion Paper No. 8922
March 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 8922
March 2015

ABSTRACT

Cultural Biases in Migration: Estimating Non-Monetary Migration Costs*

Ever since Sjaastad (1962), researchers have struggled to quantify the psychic costs 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 historical dialect dissimilarity between German counties. Conditional on geographic distance and pre-migration wage profiles, we find that migrants demand a (indexed with respect to local rents) wage premium of about 1 (1.5) percent for overcoming one standard deviation in cultural dissimilarity. The effect is driven by males and those who earn above average occupational wages before migration, more pronounced for geographically short moves, and persistent over time.

JEL Classification: D51, J61, R23

Keywords: migration costs, culture, internal migration, psychic cost

Corresponding author:

Jens Ruhose
Ifo Institute
Poschingerstr. 5
81679 Munich
Germany
E-mail: ruhose@ifo.de

* We thank George Bulman, Thiess Buettner, Robert Fairlie, David A. Jaeger, Alexander Karmann, Constantin Mang, Fabian Waldinger, Simon Wiederhold, Ludger Woessmann, Michael Wyrwich, and seminar participants at the Ifo Institute, 3rd workshop for regional economics at the Technical University of Dresden, University of Luxembourg, University of California-Santa Cruz, CUNY Graduate Center, 2014 workshop of the German Committee of Regional Economist, the 2014 meeting of the German Economic Association, IWH Halle, and Institute for Employment Research (IAB) for helpful comments and discussion.

1. Introduction

"Culture matters" has become a general wisdom in the economics literature. Indeed, recent literature shows that norms and values, such as trust, which is an important component of culture, determines economic activity and eventually growth (e.g., Algan and Cahuc 2010; Guiso *et al.* 2006, 2011; Tabellini 2010). Specifically, cultural biases hinder economic exchange across locations (Belot and Ederveen 2011; Comin et al. 2012; Felbermayr and Toubal 2010; Guiso *et al.* 2009; Mayda 2009; Spolaore and Wacziarg 2009). Since culture only changes slowly over time[1], it is also a main source of the legacy of history emphasized by economic historians (Nunn 2009). We contribute to the economics-of-culture literature by quantifying the non-pecuniary cost of migrating to a culturally different location. The key problem in quantifying this cost is to disentangle differences in culture from other (formal) institutional differences. To solve this problem, we look inside a single country, where formal institutions do hardly differ across locations, and study internal migration across locations within a country. Thereby, we exploit cultural biases between locations coming from a lack of social and economic interactions in the past.

Quantifying migration costs is important because labor immobility is associated with substantial welfare losses as it prevents the efficient allocation of labor (Clemens 2011). Welfare losses might also arise because the mobility of workers is a precondition for cluster formation and the realization of agglomeration economics (Duranton 2011). Worker mobility also leads to the rapid dissemination and cross fertilization of ideas, which eventually fuels innovation (Saxenian 1994). Disentangling cultural costs from other costs of migration is thereby important from a policy perspective because policy can influence travel cost, e.g. by providing adequate infrastructure, but it can hardly influence deeply rooted cultural biases.

Conceptually, 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 (Todaro 1969; Harris and Todaro 1970). Benefits and costs can be monetary or non-monetary; also including the non-monetary psychic migration costs of moving from a familiar to an unfamiliar surrounding (Sjaastad 1962). Examples for these costs are that migrants might have to leave family and friends or have to cope with different cultural traits and habits in the new destination. However, it is hard find direct measures of these psychic costs. For this reason, the internal migration literature typically uses simple geographic distance between the place of origin and destination as a catch-all proxy for various costs of migration (cf. Greenwood 1975).[2] That is also the reason why the more current literature that tries to quantify non-pecuniary migration costs is still rather scarce and tailored to specific populations. For example, Barrett and Mosca (2013) find that Irish male return migrants are

[1] Guiso *et al.* (2006) define culture as "those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation."

[2] Schwartz (1973, p. 1161) justifies the use of this particular catch-all proxy: "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".

more likely to have suffered from alcohol problems than those who never moved. Another paper, dealing also with internal migrants, by Dahl and Sorenson (2010) looks at skilled technical workers in Denmark. They find that these workers are more likely to accept job offers from regions close to where they grew up. Based on conditional logit regressions, the paper finds substantial psychic migration costs through some back-of-the-envelope calculations to quantify, for example, the effect of doubling the distance to home or doubling the distance to parents.

We quantify the psychic costs of migration as the wage premium that migrants demand when moving to a culturally different location. This approach is motivated by regional general equilibrium models in the tradition of Roback (1982). We assume that living in a culturally unfamiliar environment is comparable to a disamenity in the Roback model. Consequently, a potential internal migrant will move to a culturally unfamiliar environment only if she is compensated for this disamenity in the form of a higher wage and/or lower rent compared to her place of origin. For this purpose, we use administrative social security panel data to identify internal migrants in Germany. Internal migrants are defined as job switchers for whom the job switch involves the move from one county to another.[3] 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 the time, dialects were still the prevalent languages of communication, often leading to significant problems in understanding between people from different regions or even nearby towns. 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."[4]

Our findings imply that, conditional on geographic distance and quarterly pre-migration wage profiles, internal migrants demand a wage premium of about 1 percent for overcoming one standard deviation in historical dialect distance. If we index wages with

[3] Thus, we do not study commuting. We are also not looking at gains to migration in general. McKenzie *et al.* (2010) have shown that these gains are hard to retrieve from non-experimental data. Therefore, we focus solely on the internal margin, that is, looking only at individuals who have changed their place of work and place of residence and do not discuss outcomes from individuals who have changed their place of work without moving to another region. Refer to Lehmer and Ludsteck (2011) for a study about gains to interregional migration within Germany.

[4] The linguistic situation changed when social and economic exchange was intensified after the founding of the German empire in 1871. At that point, the national language (Hochdeutsch), which, until then, had been mostly reserved for written contexts, became increasingly adopted for speech also. At the same time, and considerably more so after World War II, German dialects show signs of both convergence and linguistic transfer from the national language. Obtaining explicit cultural consolidations at a very small geographic scale is thus made easier by using historical dialect data.

respect to local rents, we document an indexed wage premium of 1.5 percent. This wage premium is most likely a lower-bound estimate for internal migrants since the county of immediate origin of an internal migrant is not necessarily the place where she was born and socialized. For those cases, however, we would not expect to find a systematic correlation between wage changes and dialect distance. It could also be that migrants in culturally more distant regions are discriminated and earn lower wages. However, this would also work against our finding that migrants in culturally more distant regions earn more. The effect is driven by males and those who earn above average occupational wages, 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 is not only reflecting a non-linearity in the geographic distance effect. We also analyze those who have made multiple moves within a relatively short period 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. Our results imply that each analysis of returns to migration that do not consider these psychic costs of migration overestimates the rate of return to monetary resources allocated to migration.

We interpret our findings within a model of search and matching (see for example Postel-Vinay and Robin (2002)) where cultural barriers to migration represent a labor market friction that prevents the efficient allocation of labor. The rationale is the following: Détang-Dessendre *et al.* (2004), for example, argue that the search process of people looking for a job is not random and that individuals accept wage offers only when they are compensated for (pecuniary and non-pecuniary) migration costs. Because of cultural barriers, migrants do not consider the whole spectrum of wage offers and make suboptimal choices in terms of wage income. However, individual welfare might not suffer from this search behavior because they are finally compensated by higher cultural familiarity with the destination region.

The remainder of the paper is organized as follows. Section 2 introduces the data. Section 3 explains our estimation strategy. Section 4 shows the results. Section 5 concludes.

2. Data

2.1 Historical Dialect Distance between German Counties

Our proxy for cultural distance is based on historical dialect data from German localities. This unique source of data is derived 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.[5] One of the chief results of this project was the discovery of 66 prototypical

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

characteristics of pronunciation and grammar that Wenker and his successors isolated during an extensive evaluation process (cf. Wrede *et al.* 1927). These characteristics are most relevant for structuring the German-language area. Hand-drawn maps illustrate the distribution of each of these prototypical characteristics across German counties.[6]

Using these prototypical characteristics, Falck *et al.* (2012) construct a dialect similarity matrix across all German counties. Based on the hand-drawn maps, they map the 66 prototypical characteristics to the current boundaries of the 439 German counties (NUTS-3 level) and count for every pair of counties how many of the characteristics can be found in both dialect profiles. Thus, the dialect similarity is 0 if there is no overlap in the dialect and 66 if all characteristics can be matched (for more details, see Falck *et al.* 2012 and Lameli 2013). The authors end-up with a dialect similarity matrix of the dimension 439 x 439.

$$Dialect\ distance_{sd} = 1 - Dialect\ similarity_{sd}/66 \in [0,1] \tag{1}$$

For means of comparison with geographic distance., we convert thedialect similarity measure from a similarity matrix into a distance matrix by taking one minus the dialect similarity divided through the maximum value of 66 (Equation (1)). The resulting dialect distance matrix across all counties has a dimension of 439×439, with elements ranging between 0 (dialect identity) to 1 (maximum dialect distance). The subscript *s* indicates the sending county and d the destination county. 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 to the south of Worms. The map documents that counties closer to Worms are, on average, culturally more similar to Worms. However, if we draw a concentric circle around Worms, which we do when we control for geographic distance, there are counties on that circle that are culturally more and less similar to Worms. We exploit this variation in our empirical setup later on.

<<Figure 1 about here>>

2.2 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 is a 2 percent subsample of the universe of employees who are subject to social security in Germany. Besides gross monthly wages, the data provide information on age, gender, educational attainment, nationality, and place of work and residence.[7] Our sample period

[6] All hand-drawn maps are published online as the 'Digitaler Wenker-Atlas' (DIWA), see http://www.diwa.info or, more recently, http://www.regionalsprache.de.

[7] 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).

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 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. Another problem is that the wage data are top-coded at the social security maximum. The number of workers affected by this restriction in the full sample is of the order of 10 to 12 percent of male workers and 1 to 3 percent of female workers (Card *et al.* 2013). The literature proposes imputing the missing wage information by assuming a normal wage distribution (Dustmann *et al.* 2009; Card *et al.* 2013). However, we restrict our sample to include only persons who have moved between two quarters. We find that only about 2 percent of the movers, either one quarter before or after the move, report top-coded wages. Thus, in total, we have only about 4 percent of top-coded observations. Therefore, instead of using imputation methods, we check the robustness of our results by excluding this group and find that the results do not change (Column (1) of Table 3).[8] Finally, we restrict our sample to German citizens only because it is not clear how dialect distance should affect those born abroad.

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.[9] 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.[10]

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*). Wages are in 2010 prices, that is, we adjust wages for the national consumer price index (Federal Statistical Office, 2014). The average monthly gross wage in 2010 prices before the move is € 2,867 and increases by 3.9 percent to € 2,980 after the move.

$$wage_{ict}^{indexed} = \frac{wage_{ict}}{index_c} \tag{2}$$

[8] The social security data should only report wages until the social security maximum. However, there are a few cases in which the reported wage exceeded this amount. 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 to this omission.

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

[10] 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 only of working individuals subject to social security, the internal migration rate in our sample is slightly lower.

To account for the fact that moving to a culturally unfamiliar environment might also be capitalized in rents, we also calculate an index wage based on local rents. We use rental prices averaged over the years 2004 to 2008 as reported by the Federal Institute for Research on Building, Urban Affairs and Spatial Development, as well as by the IDN ImmoDaten GmbH. The rental prices are transformed into a price index ($index_c$) expressed in terms of the most expensive place, Munich. Thus, the index ranges between 1 for Munich and 0.35 for the county Hof in Northern Bavaria. Equation (2) gives the indexed gross wage (in 2010 prices) of individual i in county c at time t. The average monthly gross wage (relative to Munich) before the move is 5,200 and increases by 1.4 percent to 5,271 after the move (see Panel A of Table 1).

<< Table 1 about here >>

Another observation from Panel A of Table 1 is that the average wage for movers to counties farther away than the median dialect distance is higher than that of movers to counties closer than 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 "far" movers have higher skills than "close" movers. However, these skills are also reflected in pre-migration wages.

Panel B of Table 1 shows descriptive statistics for the distance data. The geographic distance definitely correlates with dialect distance.[11] The mean geographic distance is 318 km (197.6 miles) for individuals who moved to a county farther away than the median dialect distance whereas the destination county is only 76 km (47.2 miles) away for the closer-than-median mover. On average, an internal migrant moves 200 km (124.3 miles) and experiences 0.372 in cultural distance by doing so.

The selection of individuals into moving across cultural borders can be seen in Panel C of Table 1. Above-median movers are more likely to have a 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 both the below-median mover group and in the above-median mover group. Age and

[11] The correlation between dialect distance and geographic distance is 0.7697. Appendix Figure A1 provides a graphical illustration. We construct the figure by portioning the dialect distance into 20 equal sized bins (5 percentage intervals) and compute the mean geographic distance for each of the bins. The relationship between both distances shows that a higher dialect distance is associated with a higher geographic distance. One standard deviation increase in the dialect distance is associated with an increase in geographic distance by 130 km (80.8 miles). The curvature follows an s-shaped curve, with an accelerating increase in geographic distance at low levels of dialect distance and decreasing increases in geographic distance at higher levels of dialect distance. This curvature is in line with the argument that dialect distance (or cultural distance) might explain non-linearity in geographic distance. The positive correlation between geographic distance and dialect distance makes it important to control for geographic distance in the following analyses. We also control for non-linearities in geographic distance in some specifications. This can be viewed as a conservative approach since it removes some of the non-linearity that might have its origin in culture.

(potential) experience are comparable across both groups.[12] The average age of the movers is 32. The question arises whether the wage earned around age 30 is a meaningful reflection of an individual's lifetime productivity or earnings. Studies by Haider and Solon (2006), who look at the relationship of current and life-time earnings, and by Chetty *et al.* (2014), who look at the relationship of parental and child earnings, show that measures using wages at age 30 are fairly 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. Estimation Strategy

3.1 Identification

The purpose of the paper is to show whether and how much people value cultural familiarity to a county. Intuitively, the identification problem can be illustrated in Figure 1 where we plot the dialect distance for each county to the city of Worms. If we think of one exemplary circle around Worms, there are counties on that circle that are culturally closer to Worms – counties to the east, west, and north of Worms – and there are counties that are culturally farer away to Worms – counties to the south of Worms. The empirical question is whether internal migrants to the south (culturally farer away) demand higher wages than internal migrants to the north (culturally closer). If we could randomly allocate migrants to the north and the south of Worms, our treatment effect would be the mean difference between the north and the south in post-migration wages. However, the location choice is not random. For example, Bauernschuster *et al.* (2014) show that skilled and risk-loving people are more likely to cross cultural borders. We can control for the educational background of movers, however, we cannot directly control for unobserved individual characteristics, such as risk-aversion, which are important determinants in the location choice. Thus, comparing unconditional post-migration wages of migrants who have moved to culturally closer regions with wages of migrants to culturally more farer regions is confounded by these unobserved individual characteristics.

To deal with self-selection into different locations, we adopt an estimation strategy from the labor economics 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 because unobserved individual characteristics should already show up in pre-treatment wages. McKenzie *et al.* (2010) evaluate the transferability of this estimation strategy to the context of gains to migration. They analyze the performance of different estimators in identifying the gains to migration causally. The authors make use of a special arrangement between New Zealand and Tonga in which the New Zealand government granted randomly the right to

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

immigrate to Tongan applicants. After extensive surveys among applicants and non-applicants, they find evidence that applicants are positively selected along observed and unobserved characteristics. They show that in the New Zealand–Tonga case, a differences-in-differences estimator, taking the wage before migration into account, is among the better performing non-experimental estimators. However, they also show that this estimator does not recover experimental estimates, which is also true for instrumental variable estimators. However, our data and approach allows us to improve on their setup.

First, McKenzie *et al.* (2010) only have one observation for pre-treatment wages. Kratz and Brüderl (2013) point out that internal migrants in Germany are also selected based on wage growth instead of wage levels. Thus, the parallel trends assumption of the differences-in-differences estimator is clearly violated. In addition, there is a one year gap between the pre- and post-treatment wage observation, which are collected by a survey. Instead, we use administrative quarterly panel data and several lags of pre-treatment wages as our control variables. This more detailed wage profile should control more accurately for unobserved individual characteristics. The approach should also control for the possibility that internal migrants are selected based on wage growth instead of wage levels. Second, McKenzie *et al.* (2010) cite the labor economics literature by arguing that non-experimental estimators perform better when the probability for receiving the treatment is more balanced in the population (common support). This should be more the case in our exercise because we look at internal migration which has lower barriers to migration than international migration. Third, we only look at the internal margin of migration, that is, *where* to move, whereas McKenzie *et al.* (2010) look at the external margin, that is, *whether* to move at all. Conditioning on the sample of movers, all individuals are more likely to receive the same treatment, that is, to choose a familiar or unfamiliar region.

Local amenities, such as schools, transport infrastructure, health care providers, shopping alternatives, or leisure facilities, and also disamenities, such as pollution, congestion, and the like, are also capitalized in local wages and rents. However, the local amenity level should not bias our estimate as long as the difference in amenities between two counties is not correlated with dialect distance. We check this assumption by controlling for several pair-wise county characteristics in the robustness checks (Column 5 of Table 3). Among other things, we include two differences in local amenities, that is, differences in weather conditions and differences in per-capita expenditures on local amenities (expenditures on schools, theaters, baths, and public parks). None of these pair-wise controls significantly alters the coefficient on dialect distance. Specifically for expenditures on local amenities, we find that higher per-capita expenditure levels in the destination than in the source county substitute for higher post-migration (indexed) wages, however, they do not affect the coefficient on dialect distance.

There are two potential biases that work against our hypothesis of a wage premium for moving to a culturally more unfamiliar environment: First, our data does not allow us to pick up the people in the county where they have grown up or got socialized. However, if we think of a complete random distribution of people in the extreme, that is, that actually no one is attached to the county where he or she is coming from, we should not observe a significant correlation between dialect distance and (conditional) post-migration wages. Second, Grogger (2011) finds wage discrimination against African Americans in the United States based on their speech patterns. If internal migrants in Germany are also discriminated in regions that are culturally farer away from their former home county, we should observe that they earn less instead of receiving a wage premium for moving to that more unfamiliar region. Thus, discrimination would work against our hypothesis of a wage premium. In that sense, we interpret our results as lower bound estimates of the effect of cultural distance on post-migration wages. This is underlined by the fact that we can only trace out effects from people that actually move; which are only 3 percent of the working population (see above). For non-movers, it seems that migration costs are so high that they do not consider moving at all, even though there are substantial differences in wages across German regions (Buettner and Ebertz 2009).

3.2 Empirical Setup

Following the identification discussion from above, we estimate the following wage regression:

$$\begin{aligned}\log wage_{idt+1}^{indexed} &= \alpha + \beta\ dialect\ distance_{sd} + \sum_{j=1}^{4} \gamma_j \log wage_{ist-j}^{indexed} \\ &+ \lambda\ geographic\ distance_{sd} + \phi X_{it-1} + \mu_t + \epsilon_{idt+1}\end{aligned} \tag{3}$$

The log of the indexed wage (see Equation (2)) received by internal migrant i in destination county d in the quarter after the move, that is, at $t+1$, is regressed on the dialect distance between the origin county s and destination county d (see Equation (1)). We control for the (linear) geographical distance[13] between the two counties as well as gender, education (five dummies), experience (and its square), and a dummy indicating an industry change accompanying the move.[14] The quarter-year fixed effects μ_t capture all time-specific shocks.[15] Finally, ε_{idt+1} is an idiosyncratic error term. We use robust standard errors throughout the paper.[16]

[13] We deal with non-linear geographic distance in Section 4.2.
[14] An alternative for using a dummy for the industry change is to use industry fixed effects. The results do not change.
[15] In a robustness check, we also use source county fixed effects. The results do not change. Appendix Table A1 provides the results.
[16] 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 almost the same standard errors.

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 β reports the causal effect of dialect distance on the wage after the move requires that dialect distance not be correlated with unobserved individual characteristics. Guided by the discussion from above, we argue that we can control for unobserved individual characteristics to a large extent by including the last four quarterly wages before the move, $\sum_{j=1}^{4} \log wage_{ist-j}^{indexed}$.[17] The identifying assumption, under which β describes a causal effect is that, conditional on the pre-treatment wage profile and observed individual characteristics, the move to a familiar versus unfamiliar region is as good as random. Because there could be still omitted variables that bias the coefficient in one way or the other (see discussion above), we are confident that this strategy leads to a close to causal interpretation of the estimated treatment effect.

4. Results

4.1 Wage Premium for Overcoming Dialect Distance

Table 2 sets out our baseline results. The sample is restricted to the internal migrants' first move that we observe in the data.[18] Column (1) shows the unconditional association between dialect distance and post-migration log indexed wages. The association is positive and highly significant, meaning that a higher dialect distance leads to higher post-migration indexed wages. Column (2) shows the unconditional association between geographical distance and post-migration log indexed wages. The coefficient is also positive but not significant. Including both variables in Column (3) more than doubles the coefficient on the dialect distance and decreases the coefficient on geographic distance, which is significantly negative now. Thus, the small wage premium implied by the positive coefficient on geographic distance in Column (2) is completely captured and reversed by dialect distance. The negative coefficient on geographic distance can be explained by the construction of the dependent variable, which has the (index of the) rental price of the destination county in the denominator. We come back to this issue at the end of this section.

<< Table 2 about here >>

In Column (4), we add the last four quarterly pre-migration log indexed wages to control for unobserved individual characteristics that could drive unobserved self-selection into regions. 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 that neglecting pre-migration wages in the regression will lead to an upwardly biased effect of dialect distance. After controlling for pre-migration wage profiles, adding further control variables in Column (5) does not lead to a

[17] Using four quarterly pre-migration wages as controls is somewhat arbitrary. We check this specification by changing the structure of the pre-migration wage profiles in Section 4.2.
[18] We analyze multiple-time movers in Section 4.3.

statistically significant change in the coefficient on dialect distance. The coefficient decreases slightly to 0.075 and is still highly significant. Thus, a one standard deviation (about 0.2) increase in dialect distance increases the post-migration indexed wage by about 1.5 percent.

In Column (6) of Table 2, we provide an alternative specification in which we relate the log wage and control for the log rental price in the destination and the source county on the right-hand side of the regression. This specification gives us an indication whether the effect of dialect distance comes from a tradeoff in wages (nominator of the dependent variable) or in rental prices (denominator of the dependent variable). In addition, to get an impression of the importance of the tradeoff, the specification allows us to compare the effect of dialect distance to recent collective wage agreements, which we use as a benchmark. Note that we use in this model non-indexed quarterly pre-migration wages. Not surprisingly, we find that the log rental price in the destination county is a significant predictor of post-migration log wages. Thus, wages in areas with high rental rates are also relatively high. The rental price in the source county is not associated with the post-migration wage. Given rental prices in the destination and the source county, increasing dialect distance by one standard deviation increases the post-migration wage by about 1 percent.[19] This model also reveals that geographic distance is not correlated with post-migration wages when we condition on dialect distance, rental prices, and pre-migration wage profiles.[20] The result suggests that the negative coefficient on geographic distance in Columns (3) to (5) originates from a correlation between geographic distance and higher rental prices in the destination county. An explanation for this finding is that longer distance moves are associated with moves toward bigger cities. Internal migrants move on average 269 km when one of the 5 biggest cities in Germany (Berlin, Hamburg, Munich, Cologne, and Frankfurt) is the destination. For all other moves, we observe that people are only moving 179 km on average.

To benchmark our results, we compare the effect size found in Table 2 to the average increase in wages from before to after the move and to the most recent collective wage agreements. The indexed wages of internal migrants increase, on average, by about 1.4 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 107 percent of the average indexed wage gain in 2010 prices from internal migration. In terms of non-indexed wages, we observe an increase from the quarter before the move to the quarter after the move by 3.9 percent. Thus, in 2010 nationwide prices, the wage premium has to be 26 percent of the average wage gain in 2010 prices from internal migration. The effect size of 1 percent per standard deviation in non-indexed wages is sizeable when compared to

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

[20] Appendix Table A2 shows that the correlation between geographic distance and non-indexed post-migration wages is positive and highly significant. Adding pre-migration wages and rental prices in the destination and source county does not change this result. However, once we include dialect distance, the coefficient on geographic distance turns negative (and insignificant), which shows that the wage premium obtained from geographic distance is entirely due to overcoming cultural barriers.

estimated gains to internal migration in Germany, which are around 3 percent for the average mover (Kratz and Brüderl 2013; Lehmer and Ludsteck 2011). The treatment effect is also sizable when compared to the most recent (2013) collective wage agreements in Germany. For example, in the public sector, there was an agreed upon increase of 2.65 percent in nominal wages (ver.di 2013) and in manufacturing, an increase of 3.4 percent in nominal wages was negotiated (IG Metall 2013).

Figure 2 provides a graphical illustration of the relationship between dialect distance and post-migration indexed wages. The figure shows an added-variable plot where we use only the variation in the post-migration indexed wage and the dialect distance that remains after taking account of the full control set of the baseline model in Column (5) of Table 2.[21] The figure reveals an almost linear relationship between residual dialect distance and the conditional post-migration indexed wage once the dialect distance crosses the 10th percentile (first two bins). This ensures that the effect is not driven by outlier observations at the top and the bottom of the dialect distance distribution and that there are no substantial non-linearities.[22]

<< Figure 2 about here >>

4.2 Robustness Checks

In this section, we conduct a couple of robustness checks. We first check whether our results are driven by top-coded wages. As mentioned above, the number of top-coded wages among the movers is relatively low compared to the number of top-coded wages in the overall working population. Column (1) of Table 3 shows the results of omitting internal migrants who report a wage that is at or above the social security maximum at either directly after the move (*t+1)* or at some time before the move (*t-1* to *t-4*). However, top-coded wage observations do not affect the coefficient on dialect distance.

<< Table 3 about here >>

It could also be that the observed effect is driven by moves from and to large agglomerations, which might differ from other counties in terms of amenities. To check this, we exclude the five largest cities in Germany (Berlin, Hamburg, Munich, Cologne, and Frankfurt) as destination and source counties (Column (2) of Table 3). Even though these cities account for almost a quarter of the sample, the coefficient of the regression stays virtually the same. At contrast, restricting the sample of movers toward these five largest

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

[22] Nevertheless, we have also estimated models considering non-linear effects in the dialect distance by adding a squared term to the model. The effects size for the median dialect distance is equal to 1.6 percent (significant at 1 percent) increase for one standard deviation increase in dialect distance and, thus, very close to the baseline effect of 1.5 percent.

cities, dialect distance is only a weak explanatory variable (coefficient on dialect distance is 0.0364 and insignificant). Interestingly, the coefficient on geographic distance in this sample is positive and marginally significant at the 10 percent level. This indicates that moves toward the big agglomerations are more driven by economic rather than cultural motives.

In Column (3) of Table 3, we test the robustness of the effect by including dummies for moving from East to West Germany, West to East Germany, and moving within East Germany (moving within West Germany is the baseline category). This specification controls for important cultural differences between German regions. However, the coefficient on dialect distance is not affected.

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.[23] 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 could make a person more attached to that county than to the former home county (Burchardi and Hassan 2013). Given that our panel covers nine years, we restrict our analysis to those 1,815 individuals who resided and worked in the origin county for at least seven years and then moved to a different county during the last two years of our panel. The result of this procedure is shown in Column (4) of Table 3. The coefficient of dialect distance almost doubles, providing more support 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 robustness check (Column (5) of Table 3) introduces various pair-wise historical and contemporaneous controls between the counties of origin and destination that might be correlated with both migration flows and historical dialect distance. We include the log difference in slope, the historical rail distance, a dummy that is 1 when the dominant religion is not the same in both counties, the difference in share Catholics, the difference in the historical industry structure, the difference in the current industry structure, the difference in per-capita expenditures on local amenities, and weather controls such as the difference in temperature, difference in sunshine duration, and difference in precipitation.[24] None of these controls significantly changes the coefficient on dialect distance.[25]

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

[24] Data on 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, the difference in the current industry structure are taken from Falck *et al.* (2012). Climate data come from the Deutscher Wetterdienst (DWD). We use long-term averages from 1961 to 1990. We mapped all weather monitoring stations to counties and calculated

A crucial assumption is that the four quarterly pre-migration wages sufficiently capture the migrant's unobserved ability. Table 4 introduces alternative specifications regarding the structure of the pre-migration wage profile. Column (1) of Table 4 includes pre-migration wages one, two, and three years before the move. The coefficient on dialect distance substantially increases, which indicates that a yearly pre-migration wage profile does not capture unobserved individual characteristics better than the quarterly pre-migration wage profile from the baseline model. The next specification in Column (2) uses the average yearly pre-migration wages from the last three years as control variables. The coefficient on dialect distance is almost the same as in the baseline model. In the last column of Table 4, we include the quarterly pre-migration wages of the last three years prior to the move. In fact, this specification also shows a slightly larger coefficient on the dialect distance compared to the baseline model in Table 2.

<< Table 4 about here >>

The previous robustness check indicates that adding pre-migration wages beyond four quarters do not significantly change the results. In Table 5, we provide another specification check by shifting the timing of the move to two and three years prior to the move (placebo treatment). At that time, people should not know yet that they are moving two or three years later. Therefore, the dialect distance at the time of the move should not be able to predict wages in the following quarter. If we would find that the dialect distance is able to predict wages prior to the move, we would be worried that these people are on different wage growth trajectories long before they decide to actually migrate. But Table 5 shows that this is not the case. The coefficient on the dialect distance is close to 0 and insignificant. Especially the indexed wage one quarter for the placebo move is close to one, which signals high persistence in wages. This placebo treatment analysis gives more confidence that dialect distance, conditional on pre-treatment wage profiles and observable factors, is not capturing unobserved individual characteristics.

<< Table 5 about here >>

The next concern is the possibility that dialect distance captures non-linearities in geographic distance that do not have their origin in culture. Table 6 sets out several specifications that include non-linear geographic distance measures. Column (1) replicates the baseline regression for means of comparison. Column (2) includes the geographic distance of power

---

averages. We use state averages for missing county observations. Data on differences in per-capita expenditures on local amenities come from a special statistical evaluation for the year 2004 from the Federal Statistical Office. These expenditures include expenditures on schools, research, theaters, concerts, sport facilities, public parks, and baths.

[25] We also run a regression by adding the current average bilateral migration flow (2000 – 2006) to the model in Column (5) of Table 3. Even though the bilateral migration flow might be a bad control, the coefficient on dialect distance decreases only slightly to 0.0778 and is still highly significant. The coefficient on the migration flow is, as expected, highly significant negative, indicating that the current migration flow decreases indexed post-migration wages substantially.

two and three. The coefficient on dialect distance increases slightly, which indicates that dialect distance does not capture strong non-linearities in geographic distance. However, it could be that geographic distance is an insufficient distance proxy for dialect distance. Therefore, in Columns (3) and (4) of Table 6, we use the travel distance between counties by car in minutes as an alternative geographic distance measure.[26] Column (3) shows a specification in which we include the travel distance instead of the geographic distance. The coefficient of dialect distance decreases but remains significant. When we include travel distance to the power two and three in Column (4), we see that the coefficient increases slightly again compared to Column (3). However, travel time could be affected by dialect distance because it is very likely that transportation hubs and networks have developed along historical travel routes. Therefore, it is likely that some of the effect that should be attributed to dialect distance actually goes through travel distance. Nevertheless, the results of this exercise indicate that there are only minor nonlinear effects, if any, of geographic distance that are picked up by dialect distance.

<< Table 6 about here >>

Table 7 uses alternative measures of dialect distance to check the robustness of our results. To this point, we have used a metric measure of dialect distance. However, it could be that cultural space is dependent not only on gradual differences but on categorical ones. That is, the decision to move could be due to a difference between, for example, “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 on migration of categorical differences between smaller regions, we use a classification introduced by Lameli (2013) that captures the most prominent 13 dialect areas in Germany.[27] Column (1) of Table 7 sets out the results. The coefficient is positive and significant. A one standard deviation in the dialect distance by language area (1.04) leads to 1.11 percent higher post-migration indexed wages.

<< Table 7 about here >>

As the most important linguistic difference between German dialects is that between Low German (northern part of Germany) and High German (southern part), we further construct a dummy that substantiates the particular locality of the counties and tests for movements within the two larger areas of Low German and High German. Column (2) of Table 7 includes a dummy for moving from a High German county to a Low German county, a dummy for moving from a Low German county to a High German county, and a dummy for moving from a High German county to another High German county. The omitted category is moving from a Low German county to another Low German county. The results show that the effect of dialect distance remains robust when testing for the north-south distinction. We find,

[26] Data are provided by the Federal Institute for Research on Building, Urban Affairs and Spatial Development.
[27] The measure results from bootstrapped hierarchical cluster analysis, based on the measurement of linguistic similarity of German counties.

however, a slight north-south divide, indicating the relevance of a categorical conceptualization of cultural space.

4.3 Effect Heterogeneity

The question arises as to whether there is a group of individuals that is driving the baseline results. To answer this question, we spilt in Table 8 the sample by age (Panel A), gender and education (Panel B), education *x* gender (Panel C), geographic distance of the move (Panel D), whether the wage received before the move is above or below the average wage in the specific occupation (Panel E), and occupational change (Panel F).

<< Table 8 about here >>

Panel A of Table 8 stratifies the sample between young (below age 30) and older (above age 30) movers. In the first column, we look only at movers who are 30 or younger to discover whether age plays a crucial role in overcoming cultural distance, as argued by Schwartz (1973). The coefficient is large and significant for these younger movers. The coefficient on dialect distance for older movers in the second column is smaller and not significant. This indicates that our results are more driven by young movers than by older movers, even though the difference between both coefficients is not significant. Schwartz (1973) further argues that the interaction of geographic distance with age should indicate the importance of the psychic costs of migration. Therefore, we interacted geographic distance with age. In this specification (not shown), the interaction is not significant and the effect of dialect distance remains unchanged, indicating that dialect distance better captures the psychic costs of migration than does an interaction between geographic distance and age.

Panel B of Table 8 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 largely determines 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.[28] However, the difference between the groups is not large. Panel C shows that within the group of men, it is again the group of lower qualified migrants that shows a larger coefficient, but the differences are not significant. The results for women are insignificant again and the coefficient for lower qualified women is slightly higher than for higher qualified women. Panel D reveals that the effect comes mainly from shorter-distance moves, that is, moves less than 300 km from the former home county. Thus, the wage increases from moving to a culturally more distant county are not driven by long-distance moves as one might have expected.

[28] The group of low- and medium-qualified migrants consists of those 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 is comprised of those having a degree from the highest academic track or a university degree.

Panel E of Table 8 splits the sample by those who have earned a higher or lower wage compared to the average wage in their occupation before the move. The rationale for this split is that those who have earned a wage that is lower than the comparison wage in their occupation might move more for economic reasons and do not put so much emphasis on culture. However, those who have earned a higher wage than is typically achieved in their occupation might put a higher weight on culture.[29] To test this hypothesis, we compute average occupational wages at the national level (Columns (1) and (2)) and at the county-of-origin level (Columns (3) and (4)). Then, we compare the average wage of the four pre-migration periods (quarters) to the average occupational wage for the same period. Using either of the two occupational wages, we see that the baseline effect is mainly driven by internal migrants who earn above occupational wages before the move. This indicates that movers who can be expected to move for monetary reasons, that is, to receive a higher wage in the first instance, are less responsive to cultural differences. We also looked at the subsample of internal migrants who switch occupation when they move (Panel F). 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.[30] 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 hypothetical 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. Figure 3 illustrates the composition of the sample. As mentioned above, we have 567 individuals who move at least two times. 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.[31] However, only 32 of the 194 repatriates return to the same firm.

<< Figure 3 about here >>

Table 9 shows the results of the two-time mover analysis by first and second move and by timing of the move, that is, whether the migrant moved another time after or before eight quarters (two years). Panel A shows the results with repatriates included and Panel B shows the results without this group. The first move shows, independent of the timing, that the coefficients are larger than in the baseline sample. This indicates that these particular people,

[29] Complementary to that, it could also be that those who have earned above the average occupational wage have a better bargaining position. Thus, due to the higher outside option at home, they are able to negotiate a wage premium.

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

[31] *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 repatriates.

who we know are going to move again within the next nine years, value culture highly. The second move is more interesting. 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 people who made the wrong decision about where to live and work for the first move and are now willing to sacrifice a lot more money in return for a more familiar environment.

<< Table 9 about here >>

4.4 Persistence of Wage Premium

We now turn to the question of whether the initial effect directly after the move is persistent over time. We model this question by comparing the wage growth after the move between those who have moved to a more familiar versus those who have moved to a more unfamiliar region. For the United States, Borjas *et al.* 1992 show that internal migrants earn initially less than comparable non-migrants in the new destination, but that this penalty disappears within a few years. Thus, for our case, it could be that the initially wage penalty pays off for migrants in more familiar regions because they might feel more comfortable, integrate more quickly, and are more productive in the end. Therefore, our hypothesis is that migrants who move to culturally more familiar regions experience a stepper wage growth after the move than migrants to culturally more unfamiliar regions. Empirically, we look at wage growth after the first move by estimating the following regression:

$$\begin{aligned}\left[\log wage_{idt+k}^{indexed} - \log wage_{idt+1}^{indexed}\right]/k \\ = \alpha + \beta \ dialect\ distance_{sd} + \gamma \log wage_{ist+1}^{indexed} \\ + \lambda \ geographic\ distance_{sd} + \phi X_{it-1} + \mu_t + \epsilon_{idt+1}\end{aligned} \quad (4)$$

Conditional on the logged initial wage level after the move, we regress the average yearly wage growth from period $t+1$, that is, the first quarter after the move, to period $t+k$, on dialect distance. Thereby, k takes a maximum value of 32 (quarters), that is, 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, in the analysis of eight-year post-move wage growth, there are less than 700 internal migrants. All other control variables remain equivalent to the baseline model. As mentioned above, due to a "catching-up" process, we expect that migrants who moved to culturally more dissimilar counties will exhibit lower wage growth rates. Table 10 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 generally have 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 10 about here >>

5. Conclusion

In this paper, we quantify the psychic costs of migration by combining administrative social security panel data with a proxy for cultural difference that is based on historical dialect distance between German counties. Internal migrants demand a wage premium of about 1 percent for a one standard deviation increase in dialect distance. Using wage that are indexed by local rents, we find an indexed wage premium of 1.5 percent. Compared to the general wage gain associated with internal migration, as well as to general wage increases negotiated in recent collective agreements, this wage premium is economically substantive and persistent over time. We argue that the effects are lower-bound estimates, because discrimination in regions that a culturally more distant and picking up people in regions where they might not got socialized, lead to an underestimation of the true effect. Additionally, all non-migrants should have larger migration costs than those who migrate. This implies a further underestimation of the effect. Important effect heterogeneities arise: We observe that the effect is driven by males and those who earn above average occupational wages before the move, 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 is not only reflecting a non-linearity in the geographic distance effect. We also analyze those who have made multiple moves within a relatively short period and find that internal migrants who made a "wrong decision" in the first move correct this decision in the second move and are willing to sacrifice much more of their wages to return to a more familiar region. Our results imply that each analysis of returns to migration that do not consider these psychic costs of migration overestimates the rate of return to monetary resources allocated to migration.

We interpret our findings within a model of search and matching (see for example Postel-Vinay and Robin (2002)) where cultural barriers to migration represent a labor market friction that avoids the efficient allocation of labor. The rationale is the following: Détang-Dessendre *et al.* (2004), for example, argue that the search process of people looking for a job is not random and that individuals accept wage offers only when they compensate for (pecuniary and non-pecuniary) migration costs. We have demonstrated that cultural differences represent substantial migration costs. As a consequence, potential migrants might not consider wage offers from specific local labor markets because they would dislike to life there. Thus, because migrants do not consider the whole spectrum of wage offers, they make suboptimal choices. However, individual welfare might not suffer from this search behavior because they are finally compensated by higher cultural familiarity with the destination region.

Literature

Algan, Yann and Pierre Cahuc (2010), Inherited Trust and Growth. *American Economic Review*, 100(5), 2060–2092.

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.

Bauernschuster, Stefan, Oliver Falck, Stephan Heblich, Jens Suedekum, and Alfred Lameli (2014), Why are Educated and Risk-loving Persons more Mobile across Regions? *Journal of Economic Behavior & Organization*, 98, 56-69.

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

Belot, Michèle and Sjef Ederveen (2011), Cultural Barriers in Migration between OECD Countries. *Journal of Population Economics*, 25(3), 1077–1105.

Böheim, René and Mark P. Taylor (2007), From the Dark End of the Street to the Bright Side of the Road? The Wage Returns to Migration in Britain. *Labour Economics*, 14(1), 99–117.

Borjas, George J., Stephen G. Bronars, and StephanJ. Trejo (1992), Assimilation and the Earnings of Young Internal Migrants. *Review of Economics and Statistics*, 74(1), 170-175.

Buettner, Thiess and Alexander Ebertz (2009), Spatial Implications of Minimum Wages. *Jahrbücher für Nationalökonomie und Statistik*, 229(2-3), 292-312.

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. *Quarterly Journal of Economics*, 129(4), 1553-1623.

Clemens, Michael A. (2011), Economics and Emigration: Trillion-Dollar Bills on the Sidewalk? *Journal of Economic Perspectives*, 25(3), 83–106.

Comin, Diego A., Mikhail Dmitriev, and Esteban Rossi-Hansberg (2012). The Spatial Diffusion of Technology. NBER Working Paper No. 18534.

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

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

Détang-Dessendre, Cécile, Carine Drapier, and Hubert Jayet (2004), The Impact of Migration on Wages: Empirical Evidence from French Youth. *Journal of Regional Science* 44(4), 661-691.

Duranton, Gilles (2011), California Dreamin': The Feeble Case for Cluster Policies. *Review of Economic Analysis*, 3(1), 3–45.

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.

Federal Statistical Office (2014), Consumer Price Indices, https://www.destatis.de/EN/FactsFigures/Indicators/ShortTermIndicators/Prices/pre110.html.

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

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

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

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

Guiso, Luigi, Paula Sapienza, and Luigi Zingales (2011), Civic Capital as the Missing Link. In J. Benhabib, A. Bisin, & M. O. Jackson (Eds.), *Handbook of Social Economics*, 1A ed., Vol. 1, 417–480. Amsterdam: Elsevier, North-Holland.

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

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

Harris, John R. and Michael P. Todaro (1970), Migration, Unemployment and Development: A Two-Sector Analysis. *American Economic Review*, 60(1), 126–142.

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

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson (1993), Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *Quarterly Journal of Economics*, 108(3), 577-598.

Kennan, John and James R. Walker (2011), The Effect of Expected Income on Individual Migration Decisions. *Econometrica*, 79(1), 211–251.

Kratz, Fabian and Josef Brüderl (2013), Returns to Regional Migration: Causal Effect or Selection on Wage Growth? *Journal of Applied Social Science Studies*, 133(2), 227-238.

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*. Berlin, Boston: De Gruyter.

Lehmer, Florian and Johannes Ludsteck (2011), The Returns to Job Mobility and Inter-regional Migration: Evidence from Germany. *Papers in Regional Science*, 90(3), 549–571.

Mayda, Anna Maria (2009), International Migration: A Panel Data Analysis of the Determinants of Bilateral Flows. *Journal of Population Economics*, 23(4), 1249–1274.

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.

Nunn, Nathan (2009), The Importance of History for Economic Development. *Annual Review of Economics*, 1(1), 65–92.

Postel-Vinay, Fabien and Jean-Marc Robin (2002), Equilibrium Wage Dispersion with Worker and Employer Heterogeneity. *Econometrica*, 70(6), 2295-2350.

Roback, Jennifer (1982), Wages, Rents, and the Quality of Life. *Journal of Political Economy*, 90(6), 1257–1278.

Saxenian, AnnaLee (1994), *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press.

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.

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

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

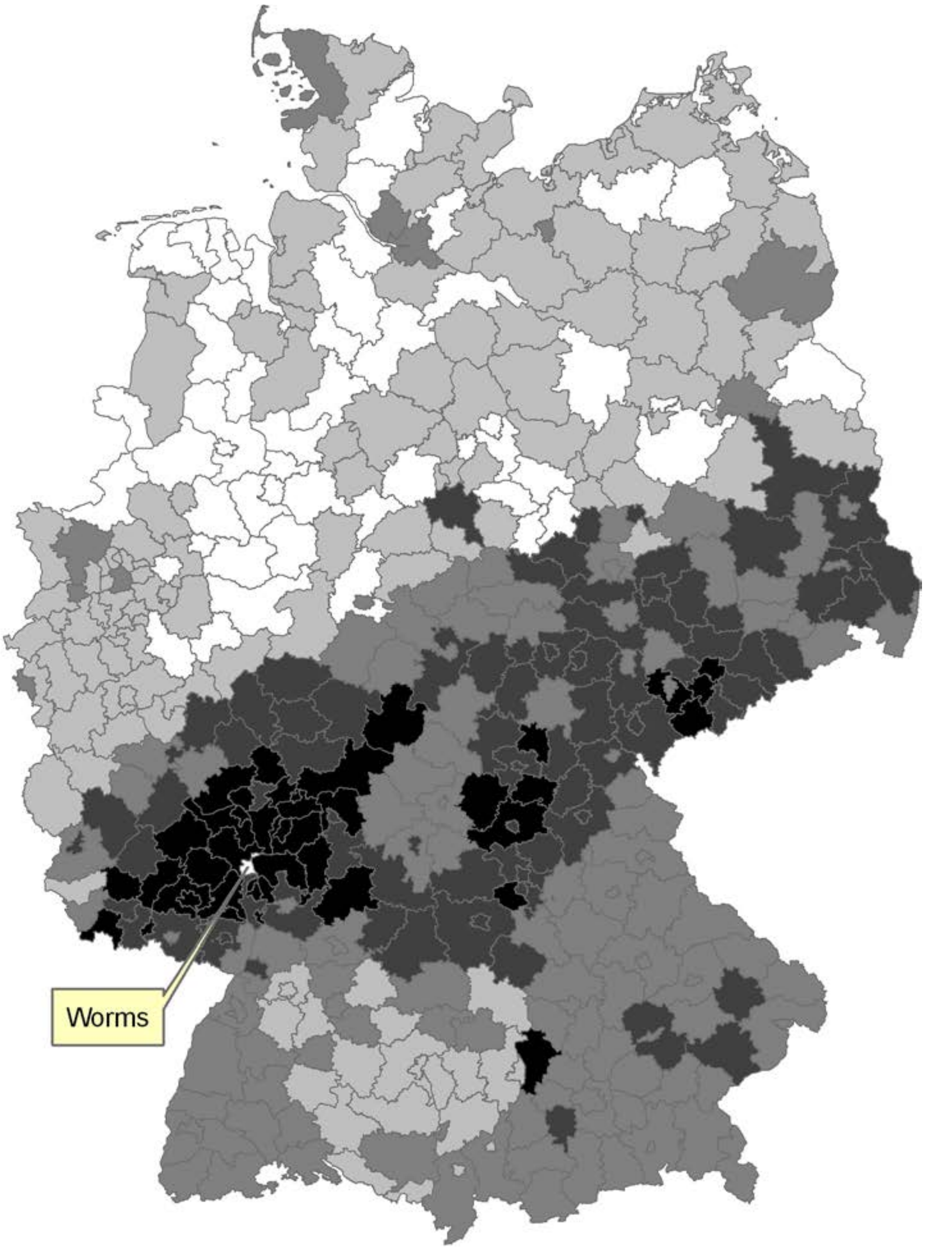
Todaro, Michael P. (1969), A Model of Labor Migration and Urban Unemployment in Less Developed Countries. *American Economic Review*, 59(1), 138–148.

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*. Marburg: Elwert.

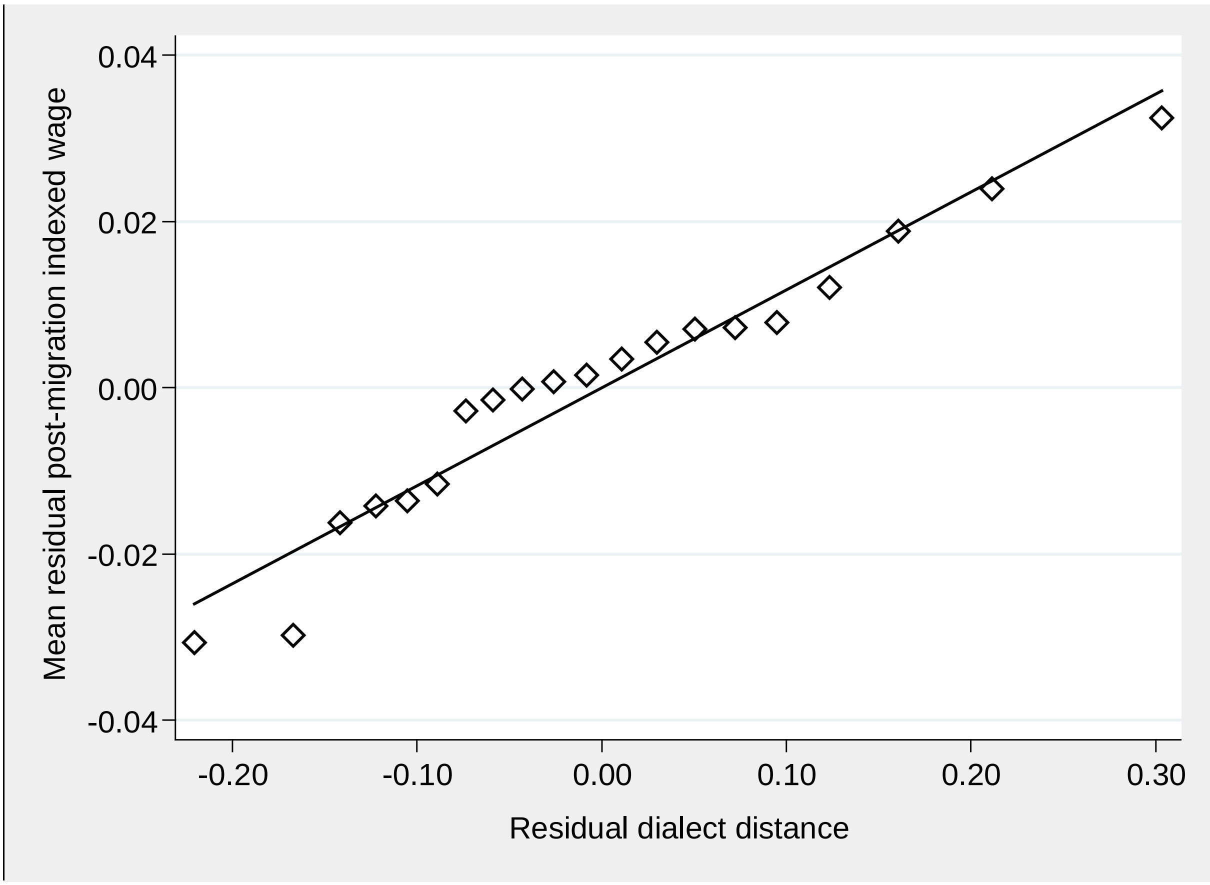
Figures and Tables

Figure 1: Dialect Distance—The Case of Worms



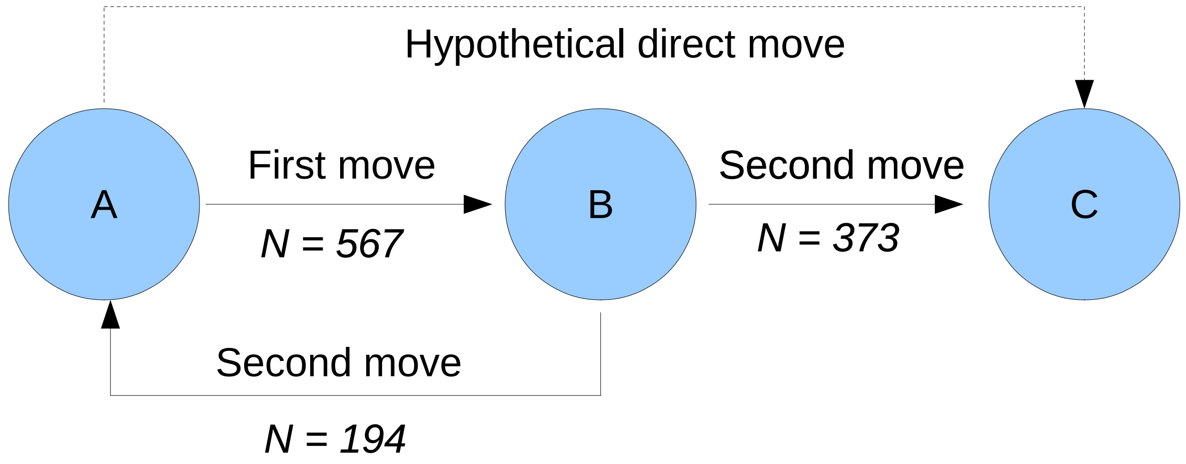
Notes: The figure shows dialect distance of all districts to the reference point Worms (20 quintiles of dialect distance). Degrees of dialect distance (from highest to lowest) are indicated by: white, grey, black.

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



Notes: The figure shows a binned scatterplot of residual post-migration indexed wages on residual dialect distance. The residual post-migration indexed wages are obtained from residuals from regressions on the geographic distance, quarter-year fixed effects, the last four quarterly pre-migration wages, education dummies, male, experience, experience squared, and an industry change dummy. The residual dialect distance is obtained from residuals from regressions on the same control set. The figure is constructed by binning dialect distance into 5-percentile point bins (so that there are 20 equal-sized bins) and computing the mean conditional post-migration indexed wage within each bin. The slope of the regression of post-migration wages, conditional on the full set of control variables is equal to 0.075 (0.029). The coefficient is significant at a 1 percent level.

Figure 3: Sample of Two-Time Movers



Notes: The figure illustrates the sample of two-time movers. In total, we have 567 internal migrants, who moved at least two times. 373 individuals moved to a different county of residence or county of work (county C) than their source county in the first move (county A). 194 individuals moved back to the same county of residence and county of work in the second move (from county B to county A). We call this group of people *repatriates*.

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 in 2010 prices (t+1)	2,980	240	2,857	248	3,098	240
	(1,291)	5,692	(1,266)	5,692	(1,303)	5,692
Wage in 2010 prices (t-1)	2,867	222	2,758	222	2,971	250
	(1,285)	5,698	(1,257)	5,698	(1,302)	5,686
Wage in 2010 prices (t-2)	2,852	223	2,744	223	2,956	225
	(1,301)	5,692	(1,270)	5,654	(1,323)	5,692
Wage in 2010 prices (t-3)	2,825	221	2,719	225	2,926	221
	(1,319)	5,716	(1,288)	5,686	(1,340)	5,716
Wage in 2010 prices (t-4)	2,806	221	2,705	227	2,903	221
	(1,329)	5,716	(1,298)	5,670	(1,351)	5,716
Indexed wage in 2010 prices (t+1)	5,271	300	5,180	300	5,358	319
	(2,402)	15,836	(2,405)	14,992	(2,397)	15,836
Indexed wage in 2010 prices (t-1)	5,200	273	5,034	273	5,358	467
	(2,371)	14,398	(2,353)	14,398	(2,378)	13,838
Indexed wage in 2010 prices (t-2)	5,174	273	5,010	273	5,331	341
	(2,398)	14,476	(2,371)	14,475	(2,414)	13,915
Indexed wage in 2010 prices (t-3)	5,129	225	4,968	225	5,282	431
	(2,432)	14,570	(2,407)	14,570	(2,446)	13,930
Indexed wage in 2010 prices (t-4)	5,100	227	4,951	227	5,243	359
	(2,444)	14,586	(2,424)	14,586	(2,455)	13,928
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
Geographic distance (km)	200	1	76	1	318	15
	(170)	818	(72)	595	(150)	818
Panel C: Individual characteristics						
Lowest and middle academic track,	0.019	0	0.021	0	0.017	0
without VET		1		1		1
Lowest and middle academic track,	0.523	0	0.572	0	0.477	0
with VET		1		1		1
Highest academic track	0.144	0	0.135	0	0.152	0
		1		1		1
University	0.300	0	0.260	0	0.339	0
		1		1		1
Education unknown	0.014	0	0.012	0	0.015	0
		1		1		1
Male	0.557	0	0.570	0	0.545	0
		1		1		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	0.578	0	0.598	0
		1		1		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. Wages are denoted in 2010 Euros by using the consumer price index from the Federal Statistical Office (2014). The dialect and geographic 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 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 vocational education and training (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)	(5)	(6)
Dialect distance	0.1691***		0.3872***	0.0952***	0.0747***	0.0486**
	(0.0249)		(0.0389)	(0.0302)	(0.0287)	(0.0234)
Geographic distance (km)		0.000018	-0.00035***	-0.00012***	-0.00013***	-0.000002
		(0.00003)	(0.00005)	(0.00004)	(0.00003)	(0.000029)
Log (indexed) wage (t-1)				0.342***	0.2608***	0.3799***
				(0.0272)	(0.0256)	(0.0273)
Log (indexed) wage (t-2)				0.1137***	0.088***	0.0984***
				(0.0334)	(0.0301)	(0.0307)
Log (indexed) wage (t-3)				-0.0134	-0.0098	0.0178
				(0.0376)	(0.0334)	(0.035)
Log (indexed) wage (t-4)				0.2000***	0.1509***	0.0897***
				(0.0293)	(0.0267)	(0.0286)
Log rental price (destination)						0.1984***
						(0.0144)
Log rental price (source)						-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	yes	yes
Observations	9,090	9,090	9,090	9,090	9,090	9,090
R^2	0.0166	0.0115	0.0225	0.4399	0.4994	0.4997

Notes: The indexed wage is the gross wage in 2010 prices divided by the index of the rental rate. Only the first observed move is considered for individuals who moved several times. Parentheses behind variables indicate x quarters before the move $(t-x)$ and one quarter after the move $(t+1)$. The omitted education category is *lowest and middle academic track, without VET*. Quarterly pre-migration wages in Column (6) are not indexed. Robust standard errors in parentheses. 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)	(4)	(5)
	Top-coded	5 largest cities	East-west dummies	7 years at origin	Pair-wise controls
Dialect distance	0.0721**	0.0741**	0.0783***	0.1295**	0.0985***
	(0.0295)	(0.0303)	(0.0291)	(0.0653)	(0.0282)
Geographic distance (km)	-0.000118***	-0.000149***	-0.000153***	-0.000158**	-0.00012***
	(0.000035)	(0.000038)	(0.000036)	(0.000078)	(0.00004)
Log indexed wage (t-1)	0.2569***	0.286***	0.2603***	0.3601***	0.2980***
	(0.0258)	(0.0297)	(0.0256)	(0.0623)	(0.0258)
Log indexed wage (t-2)	0.0945***	0.0903**	0.0877***	-0.0103	0.0806***
	(0.03)	(0.0357)	(0.0301)	(0.0783)	(0.0300)
Log indexed wage (t-3)	-0.0143	0.0304	-0.0099	0.0937	-0.0090
	(0.0335)	(0.0403)	(0.0333)	(0.0586)	(0.0329)
Log indexed wage (t-4)	0.1467***	0.1224***	0.1512***	0.0692*	0.1449***
	(0.0269)	(0.0317)	(0.0266)	(0.0363)	(0.0258)
Control variables	yes	yes	yes	yes	yes
Quarter-year fixed effects	yes	yes	yes	yes	yes
East-west fixed effects	-	-	yes	-	-
Pair-wise controls	-	-	-	-	Yes
Observations	8,727	6,946	9,090	1,815	9,090
R^2	0.4856	0.5395	0.5007	0.5327	0.5372

Notes: The indexed wage is the gross wage in 2010 prices divided by the index of the rental rate. Only the first observed move is considered for individuals who moved several times. Parentheses behind variables indicate x quarters before the move $(t - x)$ and one quarter after the move $(t + 1)$. *Control variables:* education dummies, male, experience, experience squared, industry change. Column (1) drops all movers with top-coded wages (at or above the social security maximum) at either *t+1* or *t-1* to *t-4*. Column (2) drops the five largest cities (Berlin, Hamburg, Munich, Cologne, Frankfurt) as destination and source counties. Column (3) includes fixed effects for movers from East Germany to West Germany, from West Germany to East Germany, and moving within East Germany (moving within West Germany is the baseline category). Column (4) conditions the sample on having lived at least seven years in the county of origin. Column (5) includes several pair-wise controls: log difference in slope, historical rail distance, different religion dummy, difference in share Catholics, difference in historical industry structure, difference in the current industry structure, difference in temperature, difference in sunshine duration, difference in precipitation, and difference in per-capita expenditures on local amenities (expenditures for schools, research, theaters, concerts, sport facilities, public parks, and baths.). Data on pair-wise controls come from Falck *et al.* (2012), from the Deutscher Wetterdienst (DWD) for the climate data, and from the Federal Statistical Office for expenditures on local amenities. Robust standard errors in parentheses. 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)
Dialect distance	0.1106***	0.0858**	0.0846**
	(0.0380)	(0.0372)	(0.0386)
Log indexed wage (t-4)	0.2710***		
	(0.0211)		
Log indexed wage (t-8)	0.0560**		
	(0.0231)		
Log indexed wage (t-12)	0.0851***		
	(0.0176)		
Mean log indexed wage (t-1 to t-4)		0.4161***	
		(0.0279)	
Mean log indexed wage (t-5 to t-8)		-0.0281	
		(0.03)	
Mean log indexed wage (t-9 to t-12)		0.1017***	
		(0.0204)	
Log indexed wage (t-5)			0.0331
			(0.0604)
Log indexed wage (t-6)			-0.0466
			(0.0663)
Log indexed wage (t-7)			0.0859
			(0.0665)
Log indexed wage (t-8)			-0.0914
			(0.0694)
Log indexed wage (t-9)			0.0612
			(0.0502)
Log indexed wage (t-10)			0.0659
			(0.0443)
Log indexed wage (t-11)			-0.0123
			(0.0366)
Log indexed wage (t-12)			0.0317
			(0.0311)
Control variables	yes	yes	yes
Quarter-year fixed effects	yes	yes	yes
Geographic distance	yes	yes	yes
Pre-migration log indexed wages (t-1, t-2, t-3, t-4)	-	-	yes
Observations	5,411	5,411	4,681
R^2	0.495	0.519	0.5293

Notes: The indexed wage is the gross wage in 2010 prices divided by the index of the rental rate. Only the first observed move is considered for individuals who moved several times. Parentheses behind variables indicate x quarters before the move $(t - x)$ and one quarter after the move $(t + 1)$. *Control variables:* education dummies, male, experience, experience squared, industry change. Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: Placebo Treatment

	Dependent variable: Log indexed wage (t-k)	
	(1)	(2)
	k = 8	*k = 12*
Dialect distance	-0.0256	0.0097
	(0.0174)	(0.0187)
Geographic distance (km)	0.000043*	-0.000002
	(0.000023)	(0.000025)
Log indexed wage (t-k-1)	0.7793***	0.8600***
	(0.0562)	(0.0323)
Log indexed wage (t-k-2)	0.0651	0.0244
	(0.0543)	(0.0384)
Log indexed wage (t-k-3)	0.0469	0.0517*
	(0.0343)	(0.0302)
Log indexed wage (t-k-4)	0.0282	0.0278***
	(0.0223)	(0.0105)
Control variables	yes	yes
Quarter-year fixed effects	yes	yes
Observations	5,436	3,815
R^2	0.9268	0.9476
Coefficient in baseline	0.0933**	0.1099**
	(0.0383)	(0.0474)

Notes: The indexed wage is the gross wage in 2010 prices divided by the index of the rental rate. Only the first observed move is considered for individuals who moved several times. Parentheses behind variables indicate x quarters before the move $(t-x)$ and one quarter after the move $(t+1)$. *Control variables:* education dummies, male, experience, experience squared, industry change. The *coefficient in baseline* gives the coefficient on dialect distance from the baseline regression of log indexed wage (t+1) on the full control set restricted to the respective sample. Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6: 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)
Geographic distance (km)	-0.00013***	0.00002		
	(0.00003)	(0.00018)		
Travel distance			-0.00014***	0.00029
			(0.00005)	(0.00025)
Geographic distance (km) squared x 10^{-5}		-0.0924		
		(0.064)		
Geographic 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)	(0.030)	(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)
Control variables	yes	yes	yes	yes
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: The indexed wage is the gross wage in 2010 prices divided by the index of the rental rate. Only the first observed move is considered for individuals who moved several times. Parentheses behind variables indicate x quarters before the move $(t - x)$ and one quarter after the move $(t + 1)$. *Control variables:* education dummies, male, experience, experience squared, industry change. Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: Alternative Dialect Measures

	Dependent variable: Log indexed wage (t+1)	
	(1)	(2)
Dialect distance by language area	0.0106**	
	(0.005)	
Dialect distance		0.0582*
		(0.0307)
Moving from High German to Low German		0.0847***
		(0.0148)
Moving from Low German to High German		-0.0955***
		(0.0144)
Moving from High German to High German		-0.0425***
		(0.0095)
Geographic distance (km)	-0.000104***	-0.000131***
	(0.000031)	(0.000033)
Log indexed wage (t-1)	0.2619***	0.2687***
	(0.0255)	(0.0252)
Log indexed wage (t-2)	0.0874***	0.0866***
	(0.0301)	(0.0298)
Log indexed wage (t-3)	-0.0095	-0.0077
	(0.0335)	(0.0331)
Log indexed wage (t-4)	0.151***	0.1476***
	(0.0267)	(0.0265)
Control variables	yes	yes
Quarter-year fixed effects	yes	yes
Observations	9,090	9,090
R^2	0.4992	0.5093

Notes: The indexed wage is the gross wage in 2010 prices divided by the index of the rental rate. Only the first observed move is considered for individuals who moved several times. Parentheses behind variables indicate x quarters before the move $(t-x)$ and one quarter after the move $(t+1)$. *Control variables:* education dummies, male, experience, experience squared, industry change. *Moving from High German to Low German* 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 Low German to High German* 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 High German to High German* 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 Low German to Low German*. Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 8: Effect Heterogeneities

Dependent variable: Log indexed wage (t+1)

	< 30	≥ 30
Panel A: Age	0.0930**	0.0494
	(0.0404)	(0.0407)
Observations	4,384	4,706
R^2	0.3948	0.5042

	Gender		Education	
	Men	Women	Low, medium	High
Panel B: Gender and education	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

	Men		Women	
	Low, medium	High	Low, medium	High
Panel C: Gender x education	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

	< 200 km	< 300 km	≥ 200 km	≥ 300 km
Panel D: Geographic distance	0.0984**	0.1179***	0.0534	-0.0279
	(0.0402)	(0.0356)	(0.0545)	(0.0702)
Observations	5,325	6,575	3,765	2,515
R^2	0.5155	0.5112	0.4862	0.4822

	Average wage (t-1 to t-4) compared to average national-level occupational wage (t-1 to t-4)		Average wage (t-1 to t-4) compared to average county-level occupational wage (t-1 to t-4)	
	Above	Below	Above	Below
Panel E: Compared to occupational wage	0.1141***	0.0129	0.0884**	0.0404
	(0.036)	(0.0465)	(0.0346)	(0.0496)
Observations	5,225	3,865	5,416	3,674
R^2	0.4977	0.3759	0.5153	0.3817

	Occupational information available	Occupational change	
		Switchers	Stayers
Panel F: Occupational change	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: Each panel shows the coefficient on dialect distance from a regression on the baseline model. Dependent variable in all specifications is the log indexed wage (t+1). The indexed wage is the gross wage in 2010 prices divided by the index of the rental rate. Only the first observed move is considered for individuals who moved several times. Parentheses behind variables indicate x quarters before the move $(t - x)$ and one quarter after the move $(t + 1)$. All regressions include geographic distance, four quarterly pre-migration indexed wages, education dummies, male, experience, experience squared, industry change, and quarter-year fixed effects. In Panels B and C, *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 education. Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 9: Two-Time Mover Analysis

Dependent variable: Log 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: Each panel shows the coefficient on dialect distance from a regression on the baseline model. Dependent variable in the specifications is the log indexed wage after the first or the second move, respectively. The indexed wage is the gross wage in 2010 prices divided by the index of the rental rate. In the second move, dialect and geographic distance are from the (hypothetical) direct move from county A to county C (or back to county A) as illustrated in Figure 3. All regressions include four quarterly pre-migration indexed wages, education dummies, male, experience, experience squared, industry change, and quarter-year fixed effects from the first or the second move, respectively. *Repatriates* are those individuals who move back to their county (or counties) of origin (both county of work and county of residence) in the second move. Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

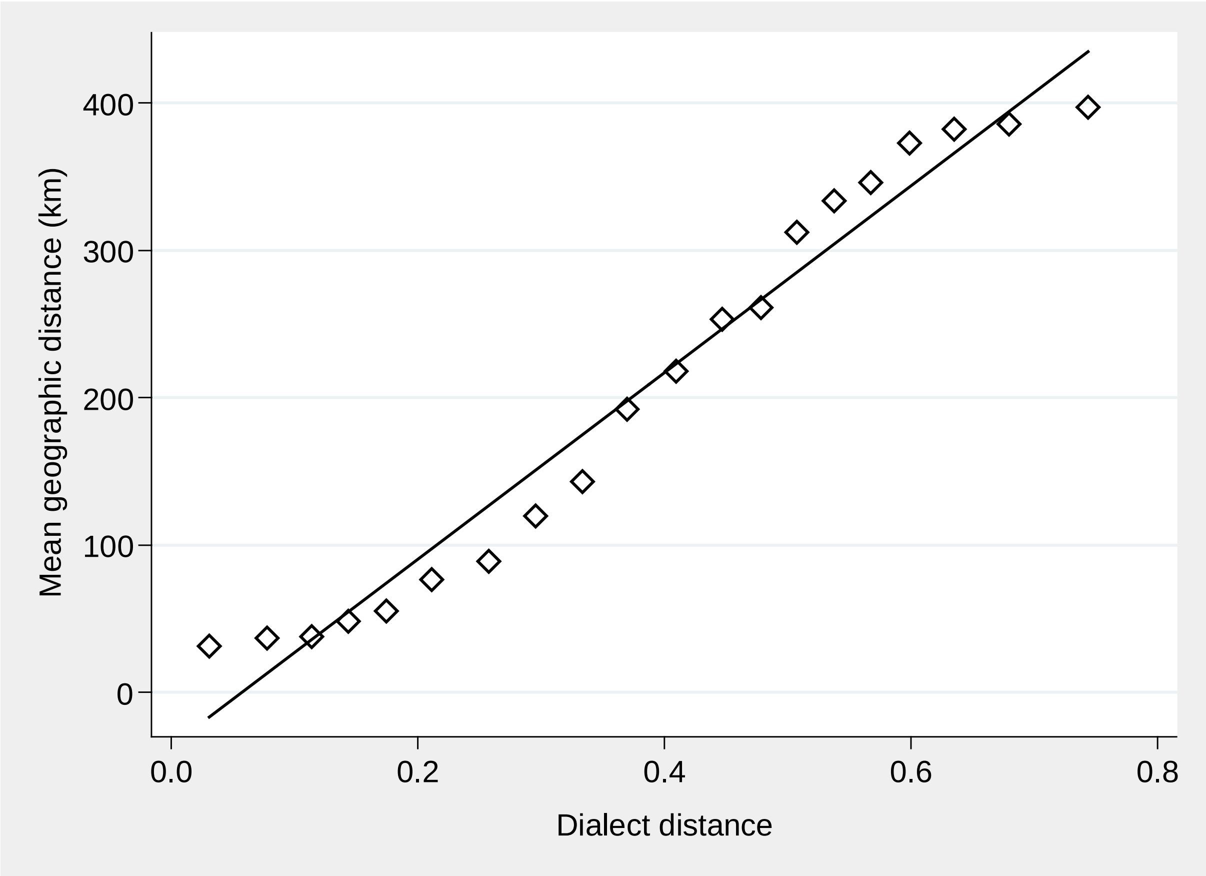
Table 10: Long-Run Effects

	Dependent variable: [log indexed wage (t+k) - log indexed wage (t+1)]/k							
	k = 4	*k = 8*	*k = 12*	*k = 16*	*k = 20*	*k = 24*	*k = 28*	*k = 32*
Dialect distance	0.0088*	0.0034	0.0022	0.0005	0.0003	0.0022	-0.0001	-0.0002
	(0.0048)	(0.0034)	(0.003)	(0.0026)	(0.0024)	(0.0024)	(0.0029)	(0.0044)
Log indexed wage (t+1)	-0.0431***	-0.0338***	-0.028***	-0.0223***	-0.0194***	-0.0166***	-0.0164***	-0.015***
	(0.0035)	(0.0022)	(0.0017)	(0.0015)	(0.0013)	(0.0013)	(0.0014)	(0.0021)
Control variables	yes	yes	yes	yes	yes	yes	yes	yes
Quarter-year fixed effects	yes	yes	yes	yes	yes	yes	yes	yes
Geographic distance	yes	yes	yes	yes	yes	yes	yes	yes
Observations	8,209	6,872	5,875	4,910	4,076	3,142	1,923	698
R^2	0.0922	0.1224	0.119	0.1317	0.1333	0.1271	0.1491	0.1627

Notes: The indexed wage is the gross wage in 2010 prices divided by the index of the rental rate. Only the first observed move is considered for individuals who moved several times. Parentheses behind variables indicate x quarters before the move $(t-x)$ and one quarter after the move $(t+1)$. *Control variables:* education dummies, male, experience, experience squared, industry change. Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix

Figure A1: Geographic Distance and Dialect Distance



Notes: The figure shows a binned scatterplot of geographic distance on dialect distance. The figure is constructed by binning dialect distance into 5-percentile point bins (so that there are 20 equal-sized bins) and computing the mean geographic distance within each bin. The slope of the regression on the microdata is 629.89 (5.436) and is significant at a 1 percent level.

Table A1: Source-county Fixed Effects

	Dependent variable: Log indexed wage (t+1)		
	(1)	(2)	(3)
Dialect distance	0.0936***	0.0956***	0.0861*
	(0.0328)	(0.0323)	(0.0483)
Geographic distance (km)	-0.000174***	-0.000146***	-0.000121**
	(0.00004)	(0.000039)	(0.000057)
Log indexed wage (t-1)	0.333***	0.2671***	0.3698***
	(0.0278)	(0.0257)	(0.0437)
Log indexed wage (t-2)	0.0959***	0.0903***	0.0935**
	(0.0317)	(0.0306)	(0.0463)
Log indexed wage (t-3)	-0.0193	-0.0092	-0.0453
	(0.036)	(0.0352)	(0.0439)
Log indexed wage (t-4)	0.151***	0.1517***	0.155***
	(0.0285)	(0.0283)	(0.0412)
Residence county fixed effects (source)	yes	-	-
Work county fixed effects (source)	-	yes	-
Residence x work county fixed effects (source)	-	-	yes
Control variables	yes	yes	yes
Quarter-year fixed effects	yes	yes	yes
Observations	9,090	9,090	9,090
R^2	0.5503	0.5284	0.719

Notes: The indexed wage is the gross wage in 2010 prices divided by the index of the rental rate. Only the first observed move is considered for individuals who moved several times. Parentheses behind variables indicate x quarters before the move $(t-x)$ and one quarter after the move $(t+1)$. *Control variables:* education dummies, male, experience, experience squared, industry change. Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A2: Geographic Distance and Non-indexed Post-migration Wages

	Dependent variable: Log wage (t+1)			
	(1)	(2)	(3)	(4)
Geographic distance (km)	0.000186***	0.00009***	0.000057***	-0.000014
	(0.000029)	(0.000019)	(0.00002)	(0.00003)
Dialect distance				0.0756***
				(0.0246)
Log wage (t-1)		0.4797***	0.4684***	0.4657***
		(0.0301)	(0.0297)	(0.0297)
Log wage (t-2)		0.1223***	0.1208***	0.1214***
		(0.0352)	(0.0348)	(0.0346)
Log wage (t-3)		0.0163	0.0103	0.0097
		(0.0398)	(0.0395)	(0.0393)
Log wage (t-4)		0.1120***	0.1209***	0.1212***
		(0.0316)	(0.0312)	(0.0312)
Log rental price (destination)			0.2095***	0.2110***
			(0.0149)	(0.0149)
Log rental price (source)			-0.0219	-0.0186
			(0.0166)	(0.0165)
Control variables	-	-	-	-
Quarter-year fixed effects	yes	yes	yes	yes
Observations	9,090	9,090	9,090	9,090
R^2	0.0169	0.5867	0.5957	0.5961

Notes: The indexed wage is the gross wage in 2010 prices divided by the index of the rental rate. Only the first observed move is considered for individuals who moved several times. Parentheses behind variables indicate x quarters before the move $(t-x)$ and one quarter after the move $(t+1)$. *Control variables:* education dummies, male, experience, experience squared, industry change. Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.