Working Papers

www.cesifo.org/wp

Same Same But Different: Dialects and Trade

Alfred Lameli
Volker Nitsch
Jens Südekum
Nikolaus Wolf

CESifo
Center for Economic Studies & Ifo Institute

Same Same But Different: Dialects and Trade

Abstract

Language is a strong and robust determinant of international trade patterns: Countries sharing a common language trade significantly more with each other than countries using different languages, holding other factors constant. In this paper, we show that this trade-promoting effect of language is likely to reflect cultural ties, rather than lower costs of communication or similar institutions. Analyzing unique data for a single-language country, Germany, we find that similarities in the local dialect between regions have a sizable and significant positive impact on intra-national trade. We interpret this finding as evidence for the effect of culture on trade.

JEL-Code: F140, F150, Z100.

Keywords: language, culture, trade costs, gravity.

Alfred Lameli
Research Centre Deutscher Sprachatlas
University of Marburg
Hermann-Jacobsohn-Weg 3
Germany – 35032 Marburg
lameli@uni-marburg.de

Volker Nitsch
Department of Law and Economics
Darmstadt University of Technology
Marktplatz 15, Residenzschloss
Germany – 64283 Darmstadt
nitsch@vwl.tu-darmstadt.de

Jens Südekum
Mercator School of Management
University of Duisburg-Essen
Lotharstrasse 65
Germany – 47057 Duisburg.
jens.suedekum@uni-due.de

Nikolaus Wolf
Institute of Economic History
Humboldt-University Berlin
Spandauer Strasse 1
Germany – 10178 Berlin
nikolaus.wolf@wiwi.hu-berlin.de

1. Introduction

Exchange transactions typically require some common understanding among business partners. For trade in goods, trading partners have to agree on the quantity and quality of the product, the terms of delivery and the conditions of payment, among other things. More notably, while standard features of an exchange are usually fixed in a contract, transactions also often involve unforeseeable situations for which common solutions then have to be found. Building on this idea, a broad range of factors has been identified that help make business negotiations easier. To facilitate communication, the fluency and use of a common language is most helpful. Also, business transactions seem to be less costly if partners share the same cultural and/or religious background or operate under similar institutional settings (such as comparable legal systems). Most of these factors have been analyzed, for good reason, in an international context. However, this empirical set-up often makes it difficult to differentiate between the effects of various country-specific features and hence spell out possible policy implications.

In this paper, we apply a different approach. Instead of analyzing cross-country differences in cultures and languages, we examine – for the first time to our knowledge – the effect of linguistic ties on trade across regions *within one* country, with a common institutional framework, which essentially speak the *same* language. Specifically, we explore economic interactions in Germany, which is de jure and de facto a single-language country (German), making use of two unique data sets. Our baseline data set contains information on intra-national trade flows across 101 German regional units over the period from 1995 through 2004. We combine this information with detailed historical data on the use of German dialects, derived from an encompassing language survey from the late 19th century.

To analyze patterns of trade, we perform a standard gravity analysis which aims to hold constant all determinants of trade other than the variable of interest, similarities in the local dialect of the German language. Similar to Nitsch and Wolf (2013), we find that the former border between East Germany and West Germany ('Iron Curtain') exhibits a remarkable persistence in shaping intra-German trade flows. More importantly, and previewing our key result, our estimates indicate that linguistic ties between regions, as measured by dialect data from about 120 years ago, have a sizable and highly significant positive impact on current trade.

Since local dialects no longer actually hinder communication within Germany today, the positive effect of dialect similarity on trade is unlikely to be rooted in (lower) transaction costs or communication frictions. Rather, we argue that our results capture the effect of culture on trade. Specifically, we note that the geography of dialects has been shaped over centuries with the evolution of regional cultural identities. As a result, dialects capture an entirety of historical interactions which have left imprints on local language structures. In our empirical analysis, we control for various alternative measures of cultural similarities across regions, including religious beliefs and the existence of historical political borders. As our findings remain robust to these perturbations, local dialects seem to provide a unique opportunity to comprehensively measure deep cultural ties across German regions which would have been difficult, if not impossible, to identify without the use of linguistic data.

1

The remainder of the paper is organized as follows. In Section 2, we briefly review the relevant literature and provide some additional background to our analysis. Section 3 describes our estimation approach and the data, followed by a presentation of our empirical results. Finally, Section 5 concludes.

2. Background

Empirical research has highlighted various cultural and institutional settings which make agreement easier between trading partners and, thus, are expected to lower the transaction costs of an exchange. The most commonly identified trade-promotion factor is the ability to communicate directly by speaking the same language. Egger and Lassmann (2012), for instance, perform a meta-analysis of more than 700 estimates of language effects on trade from about 80 studies; they find that a common language increases trade flows by, on average, about 44 percent. Mélitz (2008) examines various aspects of linguistic influences on trade and documents strong effects for a broad range of linguistic factors other than sharing the same (official) language, including indirect communication through translation, literacy and linguistic diversity.

Rauch and Trindade (2002) take a somewhat broader perspective which moves beyond plain linguistic links. Examining the importance of business and social networks, they show that ethnic ties help to facilitate trade. Specifically, they find that a greater fraction of ethnic Chinese in the total population of countries is associated with larger bilateral trade, as overseas Chinese create formal and informal associations which allow easier cooperation.[1]

A different approach to identifying cultural biases in economic exchange is applied in Guiso, Sapienza and Zingales (2009). Instead of exploring similarities in population characteristics across countries, they analyze match-specific measures which capture the level of bilateral trust between pairs of countries. In line with findings for linguistic and ethnic variables, their results indicate that cultural factors measurably influence trade; a greater perception of a country's trustworthiness leads to more business with that country, including goods trade.[2]

Accompanying attempts to identify directly factors that help overcoming informal barriers in international trade, two other empirical findings indicate that cultural ties tend to promote trade. Following McCallum (1995) and Helliwell (1998), a large body of literature documents that trade flows within countries sizably exceed cross-border trade, after controlling for standard determinants of trade such as the economic size of regions and the geographic distance between them. A plausible explanation for the strong home bias may be the ease of doing business among partners who share the same national background, even though it is still unclear what exactly explains this pattern. Another interesting empirical observation is that business procedures, especially the process of how buyers and sellers

---

[1] In similar fashion, a number of studies explore the effect of religious similarities on trade.
[2] Felbermayr and Toubal (2010) use voting behavior in the Eurovision Song Contest as a proxy for cultural proximity between countries.

match, vary by product type. Rauch (1999) shows that for homogeneous products there is only limited direct interaction between trading partners; these goods are typically traded on organized exchanges, with highly standardized transactions. For differentiated goods, in contrast, where business partners often have to agree on many specifics of a transaction (including product design), the relationship between the buyer and the seller is much more important (also see Nunn, 2007). When traders search sequentially for a business partner until a reasonable match is achieved, cultural similarities may lower search barriers (and costs) and, thus, promote trade.

In this paper, we provide a particularly strong test for the importance of cultural linkages in trade. Specifically, we contribute to the literature along two lines. First, we analyze the impact of linguistic differences when business partners speak the same language. While previous studies have examined various aspects related to the use of different languages (such as language similarities, the language composition of the population or multi-linguality), we provide the first attempt to analyze the effects of dialects of a single language, German, on trade.

Second, we focus exclusively on the patterns of trade within a country, analyzing intra-German trade flows only. Since our sample does not include cross-border transactions, we automatically exclude a wide range of other (cultural and institutional) differences between traders, typically associated with the existence of a national border.

In sum, our analysis explores patterns of trade between partners speaking the same language and sharing the same national and institutional background. In so doing, we extend existing research in the international trade literature, which had difficulties in disentangling the origins of the trade-promoting effect of a common language. Countries that share a common language are also likely to exhibit other types of linkages and similarities, which may have the potential to affect patterns of trade. In our intra-national context, in contrast, where differences in institutional and legal features are almost non-existent, the trade-promoting effect of linguistic ties is more likely to capture the 'pure' effect of culture. Finally, from a linguistic perspective, our study demonstrates the enormous impact that language has on economic outcomes, such as trade flows.

3. Data and Estimation

3.1 Data

At the heart of our empirical analysis is a newly compiled data set which comprises detailed information on pair-wise linkages between German regions. The data set combines information from two sources. First, we use data from Nitsch and Wolf (2013), which quantifies the pattern of intra-national trade in Germany for the period from 1995 through 2004. Using the geographically most disaggregated level available, the sample contains data for 101 regional units called *Verkehrsbezirke* (in short, VB).

Information on the volume of shipments between geographical units in Germany is available from two statistical sources, depending on the mode of transportation. Shipments by road are recorded by the Federal Motor Transport Authority (*Kraftfahrt-Bundesamt*); corresponding data on shipments by railway, ship and sea transport are available from the Federal Statistical Office (*Statistisches Bundesamt*). All data are reported in metric tons in 10 industry categories (*Güterabteilungen*); a detailed list of industry classifications is provided in Appendix 1.[3]

Our trade data set is large. In total, the raw data comprises more than 4 million observations (=101 exporters × 100 importers × 10 industries × 4 modes of transportation × 10 years); many of them are zero. Therefore, for our empirical analysis, we often aggregate the data. Specifically, following Nitsch and Wolf (2013), we apply unit values from the German foreign trade statistics to transform trade volumes into values, which allows aggregation of the industry series.

To this data, we match corresponding information on linguistic ties, making use of our second source of data. Falck, Heblich, Lameli and Südekum (2012) have recently computed a measure of dialect similarity in Germany based on data from a comprehensive language survey which was conducted by the linguist Georg Wenker in about 45,000 German schools between 1879 and 1888.[4] Covering the area of the entire German Empire, the survey asked pupils to read 40 German sentences, designed to reveal specific linguistic features, in their local dialect. The surviving material contains millions of phonological and grammatical observations in the form of handwritten protocols of the language characteristics recorded in the individual schools.

These raw data were integrated by Wenker and collaborators into a linguistic atlas of the German Empire (*Sprachatlas des Deutschen Reichs*). With more than 1,600 hand-drawn maps, the atlas illustrates in detail the geographical distribution of specific language characteristics across the German Empire. More importantly, Ferdinand Wrede, Wenker's successor in the linguistic atlas project, determined 66 prototypical language characteristics, related to the pronunciation of consonants and vowels as well as grammar that are most relevant for the structuring of the German language area.[5] For each of these 66 linguistic attributes, an individual map is available.[6] In our empirical implementation, we first match these 66 thematic maps with the regional classification scheme for intra-German trade flows (i.e., *Verkehrsbezirke*), using geographical information system (GIS) software. Next, we

---

[3] Nitsch and Wolf (2013) explore, in addition, a companion data set which contains intra-German shipments data in another format, covering 52 industries (*Güterhauptgruppen*) for 27 larger regional units (*Verkehrsgebiete*). However, since their empirical results turn out to be highly consistent between the two samples, we restrict our analysis to the geographically more disaggregated data set.

[4] As noted in Falck, Heblich, Lameli, and Südekum (2012), the survey was intended to be an in-depth investigation of language variation within the newly created German Empire. At the time of the survey, a standardized national language (*Hochdeutsch*) had not yet become prevalent; in fact, sometimes even people from neighboring villages were not able to properly communicate with each other.

[5] Wrede combined local extractions of variants to a dialect classification (see Wrede, Mitzka, and Martin, 1927). In contrast to more recent categorizations of the Wenker data, such as Wiesinger (1983), this classification has the advantage that it lends itself quite easily to a mathematical representation of dialects.

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

quantify, based on this mapping, the dialect of each regional unit in the form of binary variables and then construct a pair-wise dialect similarity matrix for these units.[7]

Figure 1a provides a map of the VB regional units in Germany along with an illustration of our measure of dialect similarity, using the VB unit of Augsburg as an example. Warm (cold) colors indicate a high (low) degree of similarity between a VB's local dialect of German and the dialect spoken in Augsburg as measured in the late 19th century. Not surprisingly, Augsburg tends to be linguistically closer to regions nearby than to more distant regions. However, the correlation between dialect distance and geographic distance is far from perfect. For instance, regions to the west of Augsburg, in which the Swabian dialect of German is also used, tend to share more language features with Augsburg than regions to the east. Stated differently, if a circle is drawn around our reference unit, the dialect distance to Augsburg's local language differs substantially across geographically equidistant regions. In our empirical analysis, we exploit this geographic variation in linguistic data.

As discussed at length in Falck, Heblich, Lameli, and Südekum (2012), the geography of dialects reflects the intensity of interactions between regions over previous centuries. Regional linkages such as, for instance, sharing a common religious history, having been part of the same administrative territory, or having been linked through massive migration flows can be reasonably expected to have left long-lasting imprints on local dialect structures. As a result, the dialect similarity measure also has a geographic distance component: geographically closer regions often had more intensive interaction and, therefore, developed more similar dialects and cultures.[8] Still, the intensity of pair-wise interactions is not perfectly coincident with geographic distance, as exemplified by the case of Augsburg above.

For current business transactions, the use of dialects is far less of an issue than in the 19th century, when the language data were collected. Due to wide-spread linguistic diffusion (facilitated, for instance, by national media), German individuals today communicate with each other easily in standard German, albeit often inked with different local accents. Even if dialects no longer create actual barriers to communication, however, linguistic features are still informative, as the modern use of dialects in Germany is strongly correlated with historical dialect similarity. Thus, our historical dialect data provides a unique opportunity to proxy for contemporaneous cultural differences across German regions that have evolved over centuries.[9]

3.2 Methodology

To examine the effect of dialect similarity on trade, we estimate gravity-type regressions of the following (very general) form:

---

[7] Appendix 2 describes the construction of the dialect similarity matrix in more detail.
[8] Michalopoulos (2012) provides an extensive discussion of how more interactions lead to a more similar language.
[9] The ability of linguistic measures to reveal cultural differences is widely discussed in other disciplines, including anthropology and sociology; see, for instance, Cavalli-Sforza (2000).

$$X_{ij} = \beta \, DialectSimilarity_{ij} + \Sigma_k \, \gamma^k \, TradeCosts_{ij}^k + D_i + D_j + \varepsilon_{ij}, \qquad (1)$$

where X_{ij} is the volume of shipments from region i to region j, averaged over the period from 2000 to 2004, DialectSimilarity is the log of the historical dialect similarity measure, TradeCosts is a set of other pair-specific trade cost measures, such as the log of geographic distance, D_i and D_j are comprehensive sets of fixed effects for regions i and j, respectively, and ε is the disturbance term. Following standard practice, we estimate (variants of) this equation using a Poisson pseudo-maximum likelihood estimator; standard errors are robust to clustering at the region pair level.

4. Empirical Results

4.1 Benchmark Results

Our benchmark estimation results are reported in Table 1. We proceed sequentially, beginning with the most parsimonious specification of equation (1); this specification includes, as the only pair-wise variable, our measure of interest, dialect similarity, along with exporter and importer fixed effects, so that the γ's are set to zero. As shown in the first column, for the simplest possible specification of the gravity model, the estimated coefficient of dialect similarity takes a positive and statistically highly significant value, indicating that trade between regions is indeed disproportionately large if they shared a more similar dialect of German in the past. However, since no other pair-specific features are taken into account, the linguistic variable may partly capture the impact of other factors that potentially affect the volume of bilateral trade (such as, for instance, spatial effects).

Therefore, we next include additional regressors, aiming to capture barriers to trade between VB regional units other than language, so that our specification gradually becomes more demanding. Specifically, we condition on the natural logarithm of geographic distance between regions; we also add dummy variables (i) for adjacent regions, (ii) for regions located on opposite sides of the former border between East Germany and West Germany, the 'Iron Curtain', and (iii) for regions located in different federal states of Germany. The results are tabulated in the remaining columns of Table 1. In line with the findings reported in Nitsch and Wolf (2013), distance, history and administrative borders are shown to measurably affect intra-German trade patterns; the gravity model delivers precisely estimated coefficients of reasonable sign and magnitude on all of these additional control variables so that shipments between a pair of German regions fall with distance and past and current administrative borders.[10] Moreover, with these extensions, the estimated coefficient of dialect similarity decreases in magnitude, which illustrates the empirical association between linguistic ties and other measures of pair-wise transaction costs. Still, even after holding constant those other bilateral linkages and interactions, the point estimate of β remains statistically different from zero at a conventional level of significance. Intra-German shipments are larger when regions

---

[10] Nitsch and Wolf (2013) show that these findings are insensitive to perturbations of the data, such as, for instance, accounting for intermodal trade.

shared similar language characteristics in the late 19th century, and thus are likely to have closer cultural ties even today.

For an illustration, consider (again) the region of Augsburg, for which language similarities by VB unit are mapped in Figure 1a. Our estimation results imply that Augsburg tends to ship more to geographically closer regions due to standard gravity forces.[11] In addition, trade is expected to be relatively larger for VB regional units that border Augsburg directly, for VBs located in the state of Bavaria, and for VBs located on the territory of the former West Germany. Most importantly, however, Augsburg's trade pattern is also shaped by linguistic ties. More specifically, since Augsburg is part of the Swabian dialect area, a dialect that is mainly used in the neighboring state of Baden-Wuerttemberg, Augsburg is expected to trade disproportionately with VBs in this region (such as Ulm), holding other factors constant.

Figure 1b illustrates this pattern, analyzing the residuals of a regression with all gravity variables (as reported in the final column on the extreme right of Table 1) except our dialect similarity measure. Taking account of spatial autocorrelation, the figure plots the result of a spot analysis based on the G_i^* statistic; see Getis and Ord (1996) for details. As shown, there are hot spots of trading, unexplained by economic and geographic factors, in the Swabian region (in the west of Augsburg), while there are cold spots in some parts of Bavaria (in the east of Augsburg), a pattern that mirrors linguistic ties and, consequently, cultural identities.

We check the robustness of our results along various lines. Table 2 presents some of those sensitivity analyses. We begin by varying our estimation technique. Following Nitsch and Wolf (2013), we apply negative binomial estimation; we also experiment with conventional ordinary least squares estimation. Reassuringly, our estimates turn out to be remarkably robust to these modifications. Next, we examine data for the earlier period from 1995 to 1999. Again, our key findings remain essentially unchanged. Finally, we substitute some of our regressors. Starting with the variable of interest, we replace dialect similarity with measures of dialect closeness and dialect correlation.[12] We also use road distance and travel time instead of the direct-line geographic distance. None of these perturbations has a substantial effect on our results.

4.2 Extensions

In order to further explore the plausibility of our findings, we perform a number of additional analyses. Most of these tests are designed to examine the subsample stability of our key results. Some analyses are restricted by data limitations.

---

[11] Parenthetically, we note that we control for other economic influences on trade (such as region size or industry structure) with comprehensive sets of exporter and importer fixed effects.

[12] For the construction of the alternative dialect similarity measures, we use the same raw data, but apply different techniques to combine information on region-specific language characteristics. Specifically, dialect closeness is constructed as a count measure of similar realizations of a linguistic feature, thereby following Falck, Heblich, Lameli, and Südekum (2012). Alternatively, dialect correlation is derived from the Pearson correlation between the realizations of a linguistic feature.

In a first exercise, we follow Rauch's (1999) argument that the importance of direct interactions between buyers and sellers varies by product type; ties are expected to be more relevant for differentiated goods with many specific product features than for homogeneous goods with listed prices and largely anonymous trades. However, while Rauch uses, for the classification of commodities, disaggregated information that even contains data at the five-digit SITC level, our data is at the one-digit NST/R level, a categorization that is specific for transportation purposes.[13] For this categorization, a reliable classification by the type of exchange hardly seems possible.

Table 3 tabulates the estimation results for individual goods categories at the one-digit level (*Güterabteilung*). Not surprisingly, the effect of dialect similarity on trade varies strongly across categories, for which we now analyze original trade volumes (that is, weights) instead of values. Still, it seems reassuring that the point estimates of β are consistently positive and, except for solid mineral fuels, statistically different from zero. Overall, it is not exactly clear to what extent differences in product characteristics matter for intra-national trade (when there are few organized exchanges and/or buyers may choose only among a limited group of regional suppliers).

In another exercise, we make use of the fact that our raw data are available by mode of transportation. Specifically, we hypothesize that shipments by sea are, to a large extent, determined by first nature geography such that linguistic ties may only have a limited impact on trade between regions. In contrast, alternative modes of transportation, especially transportation by road, allow for much greater geographical flexibility in shipping goods, thereby allowing the effect of dialect similarity on trade to become visible. The results in Table 4 seem to be generally consistent with such reasoning. For maritime transport, dialect similarity has no measurable effect on bilateral trade volumes, while trade patterns for goods shipped by truck and inland waterway vessel seem to be strongly influenced by linguistic ties.

We also split our sample along geographic lines. Despite the decades-long division of Germany, with very limited cross-border communication and exchange, we would expect the effect of dialect similarity on trade to hold both within and between the formerly separated territories, especially since we use historical dialect data collected long before the division. Table 5 presents the results. Confirming intuition, the estimated β coefficient is positive and significant for all groups of region pairs.[14]

Finally, it might be argued that dialect similarity captures the effects of other (omitted) factors that may shape bilateral trade intensity between regions. Following Falck, Heblich, Lameli, and Südekum (2012), we examine various measures of historical regional integration, potentially correlated with dialect similarity, which may exhibit a persistent effect on trade. Specifically, we use: a) a dummy variable that takes the value of one if a region pair was divided by an administrative border in 1815 (and zero otherwise); b) a dummy variable that

---

[13] Details are available online at
https://www.destatis.de/DE/ZahlenFakten/Wirtschaftsbereiche/TransportVerkehr/Gueterverkehr/Tabellen/NST2 007.pdf?__blob=publicationFile.
[14] We also examine the (additional) effect of dialect similarity on trade between regions located in the same federal state with inconclusive results.

takes the value of one if a region pair had different religious affiliations in 1890; c) a similarity index based on regional employment shares in 24 industries in 1925; and d) historical travel distance based on the route plan for the German railway system in 1833.

We explore the effect of each of these measures individually as well as in combination. As shown in Table 6, the historical variables seem to substantially affect contemporaneous trade patterns. Except for travel distances, the estimated coefficients of these variables take a reasonable sign and are statistically different from zero. More notably, with this extension, our key finding of a significant effect of dialect similarity on trade is not affected; the point estimate of γ remains virtually unchanged. We conclude that historical linguistic ties largely have an independent effect on interregional trade in Germany.

5. Conclusions

Detailed historical language data shows sizable differences among the dialects of the German language that are used in different parts of the country. For example, taking into account a broad range of specific language characteristics, we find almost identical dialects for Munich and Rosenheim but strong linguistic diversity between the dialects used in Paderborn and Ostwuerttemberg. Moreover, these dialect differences between different German regions are remarkably persistent over time.

In this paper, we examine the effect of dialect similarity on intra-German trade patterns. Applying a standard gravity model of trade, we find that linguistic ties strongly promote trade. Since differences in local dialects do not hinder direct communication, and since all parts of Germany share a common institutional framework, we interpret our findings as evidence for the effect of culture on trade.

References

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

Egger, Peter H. and Andrea Lassmann. 2012. "The Language Effect in International Trade: A Meta-Analysis," Economics Letters. 116 (August): 221-224.

Falck, Oliver, Stephan Heblich, Alfred Lameli, and Jens Südekum. 2012. "Dialects, Cultural Identity, and Economic Exchange," Journal of Urban Economics. 72 (September-November): 225-239.

Felbermayr, Gabriel J. and Farid Toubal. 2010. "Cultural Proximity and Trade," European Economic Review. 54 (February): 279-293.

Getis, Arthur and J. Keith Ord. 1996. "Local Spatial Statistics: An Overview," in Paul Longley and Michael Batty (Eds.) Spatial Analysis: Modelling in a GIS Environment. New York: John Wiley, 261-277.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2009. "Cultural Biases in Economic Exchange?" Quarterly Journal of Economics. 124 (August): 1095-1131.

Helliwell, John F. 1998. How Much Do National Borders Matter? Washington, DC: Brookings Institution Press.

McCallum, John. 1995. "National Borders Matter: Canada-US Regional Trade Patterns," American Economic Review. 85 (June): 615-623.

Mélitz, Jacques. 2008. "Language and Foreign Trade," European Economic Review. 52 (May): 667-699.

Michalopoulos, Stelios. 2012. "The Origins of Ethnolinguistic Diversity," American Economic Review. 102 (June): 1508-1539.

Nitsch, Volker and Nikolaus Wolf. 2013. "Tear Down this Wall: On the Persistence of Borders in Trade," Canadian Journal of Economics. 46 (February): 154-179.

Nunn, Nathan. 2007. "Relationship-Specificity, Incomplete Contracts and the Pattern of Trade," Quarterly Journal of Economics. 122 (May): 569-600.

Rauch, James E. 1999. "Networks versus Markets in International Trade," Journal of International Economics. 48 (June): 7-35.

Rauch, James E. and Vitor Trindade. 2002. "Ethnic Chinese Networks in International Trade," Review of Economics and Statistics. 84 (February): 116-130.

Wiesinger, Peter. 1983. "Die Einteilung der deutschen Dialekte," in Werner Besch, Ulrich Knoop, Wolfgang Putschke, and Herbert E. Wiegand (Eds.) Dialektologie. Berlin, New York: de Gruyter, 807–900.

Wrede, Ferdinand, Walther Mitzka, and Bernhard Martin. 1927. Deutscher Sprachatlas auf Grund des von Georg Wenker begründeten Sprachatlas des Deutschen Reichs. Marburg/Lahn: Elwert.

Figure 1: Dialect Similarity of VB Regional Units (Relative to Augsburg)
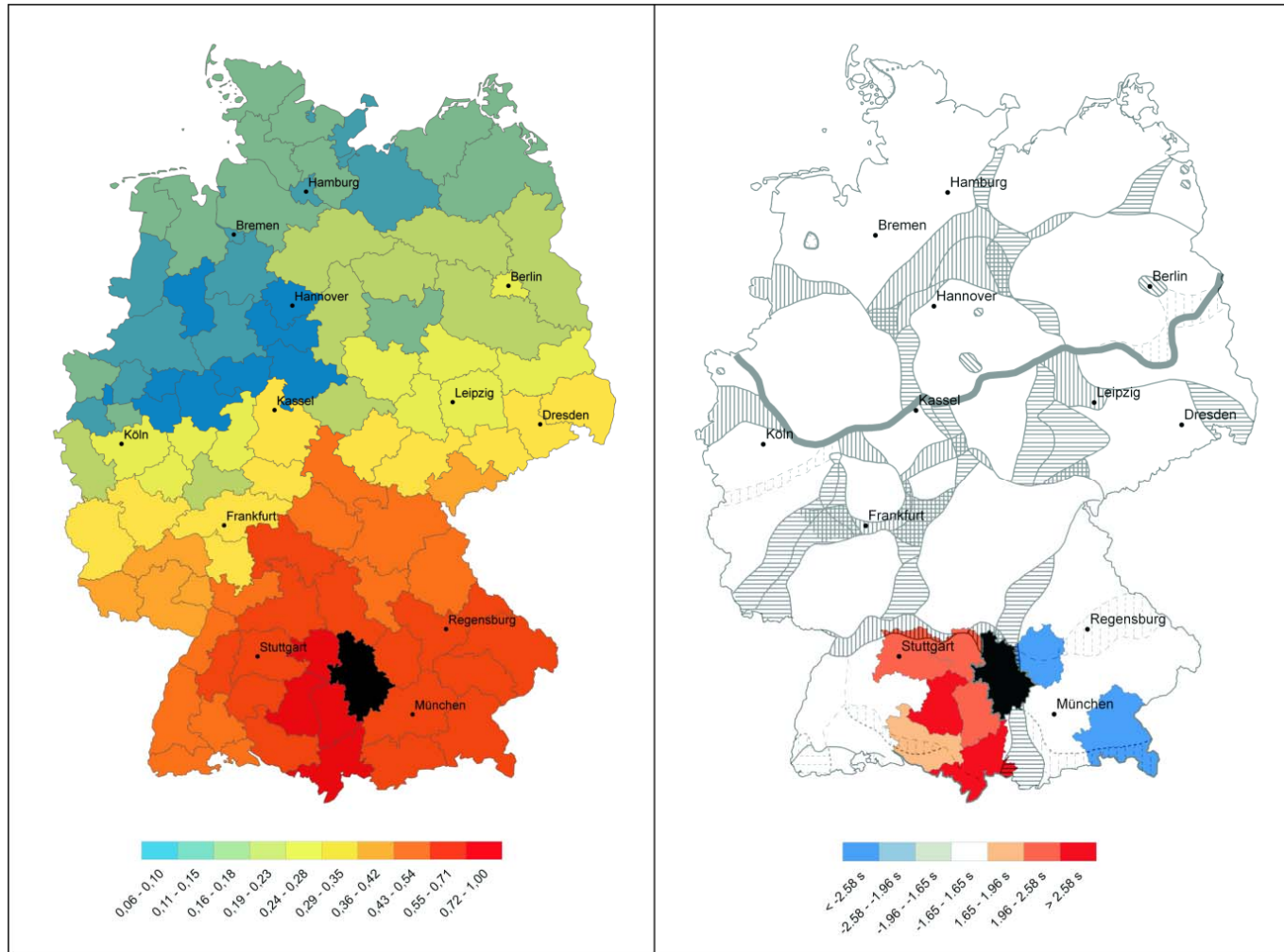
Notes: The map on the left illustrates the similarity of the linguistic dialect in VB regional units to Augsburg (black polygon), with the degree of dialect similarity decreasing from red, yellow, green, to blue. The class width follows 10 natural breaks. In the map on the right, Augsburg's trading spots are highlighted according to intervals of the standard deviation of the residual from the benchmark regression (distance, common border, E-W border, state border, fixed effects).

Table 1: Baseline Results, 2000-2004

Log Dialect Similarity	1.591	0.390	0.370	0.336
	(0.036)	(0.038)	(0.033)	(0.029)
Log Air Distance		-1.048	-0.753	-0.634
		(0.031)	(0.040)	(0.037)
Common Border			0.754	0.668
			(0.044)	(0.038)
E-W Border				-0.446
				(0.025)
State Border				-0.377
				(0.029)
R^2	0.61	0.75	0.80	0.84

Notes: Poisson estimation. Dependent variable is the value of shipments from region i to region j, aggregated from shipment volume at one-digit goods classification level. Standard errors robust to clustering at region pair level are reported in parentheses. All regressions include exporter and importer fixed effects. Number of observations is 9,700.

Table 2: Robustness Checks

Estimation Method	Negative Binomial	OLS	OLS	Poisson	Poisson	Poisson	Poisson	Poisson
Period	2000-04	2000-04	2000-04	1995-99	2000-04	2000-04	2000-04	2000-04
Dependent Variable	Shipm	Log Shipm	Log (Shipm +100)	Shipm	Shipm	Shipm	Shipm	Shipm
Log Dialect Similarity	0.336 (0.029)	0.253 (0.023)	0.289 (0.026)	0.345 (0.029)			0.315 (0.028)	0.273 (0.027)
Log Air Distance	-0.634 (0.037)	-0.962 (0.031)	-1.000 (0.035)	-0.645 (0.033)	-0.627 (0.037)	-0.658 (0.037)		
Common Border	0.668 (0.038)	0.587 (0.038)	0.515 (0.042)	0.754 (0.036)	0.664 (0.037)	0.678 (0.038)	0.657 (0.037)	0.678 (0.035)
E-W Border	-0.446 (0.025)	-0.418 (0.023)	-0.396 (0.024)	-0.553 (0.027)	-0.448 (0.025)	-0.452 (0.026)	-0.437 (0.025)	-0.416 (0.024)
State Border	-0.377 (0.029)	-0.386 (0.029)	-0.380 (0.031)	-0.400 (0.029)	-0.365 (0.029)	-0.380 (0.030)	-0.363 (0.029)	-0.357 (0.028)
Log Dialect Closeness					0.705 (0.057)			
Log Dialect Correlation						0.332 (0.030)		
Log Road Distance							-0.658 (0.034)	
Log Travel Time								-0.796 (0.037)
Number Observations	9,700	9,608	9,700	9,700	9,702	9,702	9,700	9,700
R^2		0.75	0.76	0.86	0.84	0.84	0.85	0.85

Notes: Shipm is the value of shipments from region i to region j, aggregated from shipment volume at one-digit goods classification level. Standard errors robust to clustering at region pair level are reported in parentheses. All regressions include exporter and importer fixed effects.

Table 3: Results by Goods Category (NST/R)

Goods Category	0	1	2	3	4	5	6	7	8	9
Log Dialect Similarity	0.446	0.451	0.221	1.051	0.510	0.109	0.881	0.535	0.358	0.290
	(0.043)	(0.032)	(0.233)	(0.119)	(0.108)	(0.064)	(0.058)	(0.139)	(0.052)	(0.028)
Log Air Distance	-0.709	-0.749	-1.050	-0.957	-1.058	-0.832	-0.862	-0.610	-0.542	-0.612
	(0.053)	(0.036)	(0.200)	(0.084)	(0.145)	(0.076)	(0.053)	(0.108)	(0.051)	(0.035)
Common Border	0.841	0.668	0.797	0.507	0.199	0.146	1.294	0.354	0.865	0.650
	(0.057)	(0.044)	(0.216)	(0.100)	(0.147)	(0.107)	(0.071)	(0.162)	(0.068)	(0.039)
E-W Border	-0.212	-0.303	-1.310	-0.384	-0.564	-0.092	-0.398	-0.096	-0.257	-0.416
	(0.048)	(0.030)	(0.325)	(0.148)	(0.189)	(0.071)	(0.052)	(0.211)	(0.046)	(0.025)
State Border	-0.535	-0.499	0.049	-0.674	-0.360	-0.182	-0.456	-0.890	-0.323	-0.347
	(0.045)	(0.036)	(0.246)	(0.084)	(0.094)	(0.088)	(0.048)	(0.175)	(0.048)	(0.030)
Number Zero Trade Observations	2,496	1,751	8,165	6,284	6,183	3,342	2,060	7,501	2,480	198
R^2	0.67	0.78	0.73	0.74	0.83	0.79	0.81	0.43	0.74	0.83

Notes: Poisson estimation. Dependent variable is the volume of shipments from region i to region j. Standard errors robust to clustering at region pair level are reported in parentheses. All regressions include exporter and importer fixed effects. Number of observations is 9,700 (except for NST/R goods category 2, solid mineral fuels, for which the number of observations is 9,303). See Appendix 1 for a brief description of NST/R goods categories.

Table 4: Results by Mode of Transportation

Mode of Transportation	Railway	Road	Ship	Sea
Log Dialect Similarity	-0.039	0.466	0.619	-0.607
	(0.083)	(0.028)	(0.130)	(1.255)
Log Air Distance	-0.748	-0.710	-0.897	0.016
	(0.082)	(0.037)	(0.122)	(0.449)
Common Border	0.437	0.962	0.067	-0.802
	(0.124)	(0.039)	(0.183)	(0.460)
E-W Border	-0.627	-0.358	-0.631	-0.993
	(0.097)	(0.025)	(0.173)	(0.346)
State Border	-0.303	-0.438	0.008	0.866
	(0.116)	(0.028)	(0.137)	(0.456)
Number Observations	82,044	96,314	29,391	3,210
Number Zero Trade Observations	73,834	45,292	91,545	96,380
R^2	0.81	0.75	0.68	0.85

Notes: Poisson estimation. Dependent variable is the volume of shipments from region i to region j. Standard errors robust to clustering at region pair level are reported in parentheses. All regressions include exporter-goods category and importer-goods category fixed effects. Total number of observations is 97,020.

Table 5: Results by Area

Area	West	East	E-W & W-E
Log Dialect Similarity	0.345 (0.034)	0.597 (0.083)	0.202 (0.043)
Log Air Distance	-0.581 (0.040)	-1.013 (0.070)	-1.109 (0.071)
Common Border	0.706 (0.041)	0.158 (0.090)	0.603 (0.080)
State Border	-0.395 (0.033)	-0.439 (0.053)	
Number Observations	5,850	462	3,388
Number Zero Trade Observations	74	0	0
R^2	0.85	0.89	0.81

Notes: Poisson estimation. Dependent variable is the value of shipments from region i to region j, aggregated from shipment volume at one-digit goods classification level. Standard errors robust to clustering at region pair level are reported in parentheses. All regressions include exporter and importer fixed effects.

Table 6: Historical Control Measures

Log Dialect Similarity	0.330	0.335	0.327	0.334	0.321
	(0.029)	(0.028)	(0.028)	(0.029)	(0.027)
Log Air Distance	-0.632	-0.620	-0.620	-0.630	-0.604
	(0.037)	(0.036)	(0.036)	(0.037)	(0.036)
Common Border	0.664	0.672	0.655	0.667	0.654
	(0.038)	(0.037)	(0.037)	(0.038)	(0.037)
E-W Border	-0.448	-0.439	-0.263	-0.444	-0.254
	(0.025)	(0.025)	(0.031)	(0.025)	(0.031)
State Border	-0.315	-0.371	-0.341	-0.377	-0.274
	(0.032)	(0.029)	(0.029)	(0.029)	(0.032)
Historical Political Border	-0.111				-0.109
	(0.032)				(0.032)
Religious Border		-0.128			-0.128
		(0.019)			(0.018)
Historical Diff. in Industrial Structures			-1.101		-1.129
			(0.115)		(0.113)
Historical Travel Distance				-0.009	0.000
				(0.006)	(0.006)
R^2	0.84	0.85	0.85	0.84	0.85

Notes: Poisson estimation. Dependent variable is the value of shipments from region i to region j, aggregated from shipment volume at one-digit goods classification level. Standard errors robust to clustering at region pair level are reported in parentheses. All regressions include exporter and importer fixed effects. Number of observations is 9,700.

Appendix 1: Broad Goods Categories (Güterabteilungen)

0   Agricultural products and live animals

1   Foodstuffs and animal fodder

2   Solid mineral fuels

3   Petroleum products

4   Ores and metal waste

5   Metal products

6   Crude and manufactured minerals, building material

7   Fertilizers

8   Chemicals

9   Machinery, transport equipment, manufactured articles and miscellaneous articles

Appendix 2: Construction of the Dialect Similarity Matrix

We illustrate the construction of the dialect similarity matrix with an example. One of the 66 prototypical language characteristics is the realization of the initial consonant in the German word for *pound*. In the eastern parts of Germany, *pound* is mostly pronounced as "Fund", in the northern areas as "Pund", and in the southern parts as "Pfund". This difference in pronunciation leads to the following binary code: "Fund" = {1 0 0}; "Pund" = {0 1 0}; "Pfund' = {0 0 1}. The assignment of the code to each of the 101 German regional units (*Verkehrsbezirke* or, in short, VB) is unambiguous if a specific language characteristic applies to the entire area of the VB. If more than one realization of a linguistic feature is used within a VB, the most frequent variant is considered to be representative (based on a much finer spatial grid of German NUTS 3 regions). In addition, we perform linguistic plausibility tests and cross-checks with the underlying raw data to ensure the quality of these assignments.

For each of the 66 language characteristics, the number of different realizations ranges from 2 to 18. Based on these individual language specifics, we construct K=383 binary variables which represent a linguistic 'finger print' of the dialect that is spoken in a German region in the late 19th century. Formally, the historical dialect of region *r* is represented by a vector $\mathbf{i}^r = \{i_1^r, i_2^r, \cdots, i_K^r\}$ of length K=383, where each element of the vector is a binary variable [0,1]. Overlapping the vectors for any two regions, we construct a matrix of Euclidean distances between dialects spoken in different regions of Germany. Following Falck, Heblich, Lameli, and Südekum (2012), we convert this matrix into a dialect similarity measure.